



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2000

Die 'Standards' im Vergleich mit weiteren Regelwerken zur Qualität fachlicher Leistungserstellung

Beywl, Wolfgang ; Widmer, Thomas

Abstract: Evaluation verfügt bislang in Deutschland und Österreich, vielleicht etwas weniger ausgeprägt gilt dies auch für die Schweiz, über eine schmale Professionalisierungsbasis. Es gibt z.B. keinen Ausbildungsschwerpunkt an einer Hochschule, keine Fachzeitschrift, kein verbreitetes Standardwerk, keinen Ansatz eines Berufsbildes Evaluatorin. Methodologische, methodische und ethische Fragen der Evaluation sind im deutschsprachigen Raum bislang wenig diskutiert.

DOI: https://doi.org/10.1007/978-3-322-93234-1_9

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-157844>

Book Section

Accepted Version

Originally published at:

Beywl, Wolfgang; Widmer, Thomas (2000). Die 'Standards' im Vergleich mit weiteren Regelwerken zur Qualität fachlicher Leistungserstellung. In: Sanders, James; Beywl, Wolfgang; Widmer, Thomas. Handbuch der Evaluationsstandards: die Standards des "Joint Committee on Standards for Educational Evaluation". Wiesbaden: VS Verlag für Sozialwissenschaften, 259-295.

DOI: https://doi.org/10.1007/978-3-322-93234-1_9

Die ‚Standards‘ im Vergleich mit weiteren Regelwerken zur Qualität fachlicher Leistungserstellung

Wolfgang Beywl und Thomas Widmer

I. Einleitung

Evaluation verfügt bislang in Deutschland und Österreich, vielleicht etwas weniger ausgeprägt gilt dies auch für die Schweiz, über eine schmale Professionalisierungsbasis. Es gibt z.B. keinen Ausbildungsschwerpunkt an einer Hochschule, keine Fachzeitschrift, kein verbreitetes Standardwerk⁴³, keinen Ansatz eines Berufsbildes Evaluators. Methodologische, methodische und ethische Fragen der Evaluation sind im deutschsprachigen Raum bislang wenig diskutiert.

Im Unterschied zu Nordamerika (oder auch Australien), wo Evaluation über Jahrzehnte als Disziplin gewachsen ist, an deren Reifepunkt sich die Profession Regelwerke gegeben hat, wird dies im deutschsprachigen Raum in umgekehrter Reihenfolge verlaufen: Die Regelwerke stehen hier chronologisch am Anfang einer professionellen Entwicklung und werden diese vermutlich stark mitprägen.

Für die im deutschsprachigen Raum zu leistende Entwicklungs- und Adaptionenarbeit von Regelwerken für die Evaluation hat dies auch Vorteile: Man kann dabei auf die vorliegenden evaluationsspezifischen Regelwerke vorrangig aus den USA zurückgreifen, darunter insbesondere auf die Evaluationsstandards des ‚Joint Committee on Standards for Educational Evaluation‘, (nachfolgend kurz ‚Standards‘ genannt), zum anderen auf nicht (oder lediglich partiell) auf Evaluation bezogene Regelwerke sowohl englischsprachiger als auch deutschsprachiger Fachzusammenschlüsse. Während amerikanische Ethikrichtlinien oft schon in den siebziger Jahren vorgelegt wurden, sind die deutschsprachigen überwiegend in den neunziger Jahren entstanden. Diese bieten zahlreiche Anknüpfungspunkte für rechtliche, ethische und kulturelle Anpassungen, die bei der Übertragung evaluationsspezifischer Standards in deutschsprachige Länder zu leisten sind.⁴⁴

Ein zentrales Anliegen dieses Kapitels besteht darin, die Optionen zu verdeutlichen, die bei der Abfassung von Regelwerken für Evaluationen und Evaluatoren bestehen. Das Beispiel der ‚Standards‘ ist zwar aus unserer Sicht vorbildlich und bietet weite Anwendungsfelder (siehe das vorausgegangene Kapitel „Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder“), doch sollen

43 Einen ersten Schritt dahin markiert das „Lehrbuch Evaluation“ von Heinrich Wottawa und Heike Thierau, 1998 in vollständig überarbeiteter Neuauflage erschienen.

44 So erfolgte die Erarbeitung der durch die Schweizerische Evaluationsgesellschaft (SEVAL) empfohlenen Evaluationsstandards (Widmer/Landert/Bachmann 1999) als Adaption der Program Evaluation Standards (Joint Committee 1994).

mögliche alternative Schwerpunktsetzungen verdeutlicht werden, so daß Entscheidungen auf breiter disziplinärer und kultureller Basis getroffen werden können.

Zu diesem Zweck gibt dieses Kapitel dem Leser einen Überblick über Qualitätsnormen der Evaluation oder verwandter Gebiete, welche die Standards ergänzen, mit ihnen in Konkurrenz stehen oder sie ersetzen könnten. Dabei kann keine vollständige Übersicht angeboten werden. Zu groß ist dafür die Zahl bestehender Regelwerke. Besonders bei denjenigen außerhalb der Evaluation mußten wir uns auf eine exemplarische Auswahl beschränken. Hingegen haben wir uns darum bemüht, bestehende Regelwerke der Evaluation im engeren Sinne möglichst breit abzudecken.

Folgende Dimensionen stehen im Mittelpunkt des Vergleiches:⁴⁵

1. Inhaltlicher Schwerpunkt – Was soll vorrangig durch das jeweilige Normenset reguliert werden?
2. Qualitätsträger – Wer oder was ist Träger der Qualitätsmerkmale? Sind es fachliche Leistungsprozesse oder die Fachpersonen?
3. Regelsetzer – Welche Organisation oder welches Gremium stellt das Regelwerk auf?
4. Sanktionsgehalt – Legen die Regelwerke Selbstverpflichtung auf freiwilliger Basis nahe oder sehen sie implizite oder explizite Sanktionsmöglichkeiten bei Verstößen oder Nichtbeachtung vor?
5. Umfang – Wie umfangreich sind die Regelwerke? Wie weit sind sie ausgearbeitet, dokumentiert, kommentiert und spezifiziert?

Die nachfolgenden Ausführungen sind gemäß der Übersicht in Tabelle 1 strukturiert. Darin werden zwei Ordnungsdimensionen eingeführt. Zum ersten wird danach unterschieden, ob die Regelwerke auf den Leistungsprozeß Bezug nehmen oder ob sie sich auf die Fachperson beziehen, welche Leistungen erstellt. Das zweite Merkmal unterscheidet danach, ob die Regelwerke explizit auf Evaluationen ausgerichtet sind, oder ob sie sich (auch) mit benachbarten Professionen befassen. Demzufolge ergeben sich vier Konstellationen, die in den nachfolgenden Abschnitten nacheinander thematisiert werden. In der Tabelle werden die jeweils diskutierten Regelwerke in Kurzform genannt. In den folgenden Abschnitten finden sich Erläuterungen zum einzelnen Regelwerk. Dort werden die verschiedenen Regelwerke auch mit den ‚Standards‘ verglichen. Dies gilt insbesondere bei jener Gruppe von Regelwerken, die sich – wie die ‚Standards‘ – auf evaluationsspezifische Leistungsprozesse beziehen.

Im sechsten Abschnitt werden die Ergebnisse des Vergleichs aus den vorangegangenen Überlegungen zusammengefaßt. Die Literaturhinweise zur zitierten Sekundärliteratur sind im allgemeinen Literaturverzeichnis am Schluß dieser Publikation enthalten – die Angaben zu den Regelwerken selbst finden sich direkt im Text. Verweise auf Internet-Seiten, auf denen die Regelwerke mit Stichtag 1. Mai 1999 zur Verfügung gehalten werden, sind in den Übersichtstabellen im Anhang zu diesem Kapitel enthalten.

45 Siehe auch die im Anhang zu diesem Kapitel abgedruckten Tabellen 2-5, S. 291ff.

Tabelle 1: Übersicht über die diskutierten Regelwerke nach Qualitätsträger und Evaluationsbezug

	Qualitätsträger	
	<i>Leistungsprozeß</i> (Evaluation, Supervision, Umfrageforschung ...)	<i>Fachperson</i> (Evaluatorin, Psychologe, Supervisor)
<i>evaluationsunspezifisch</i>	1. CDEP-Standards 2. DFG-Kriterien 3. DGSv-Standards 4. dv-Empfehlungen 5. GA-Standards 6. SMS-Normen	1. AERA-Standards 2. APSA-Guide 3. CMC-Kodex 4. DBSH-Prinzipien 5. DGPs/BDP Richtlinien 6. DGS/BDS-Kodex 7. DGSv-Berufsordnung 8. DGVt-Richtlinien
<i>evaluationsspezifisch</i>	1. AR-Normen 2. BAG-Leitfaden 3. ERS-Standards 4. MEANS-Handbuch 5. Ontario-Principles 6. PUMA-Guidelines 7. SWR-Prinzipien	1. AEA-Principles 2. AES-Guidelines 3. CES-Guidelines

II. Evaluationsunspezifische Regelwerke, die auf Leistungsprozesse gerichtet sind

1. Standards für pädagogisches und psychologisches Testen. Ausgearbeitet vom Committee to Develop Standards for Educational and Psychological Testing; deutsche Fassung, herausgegeben von Hartmut Häcker, Detlef Leutner und Manfred Amelang; veröffentlicht im Supplementum 1, 1998 der Diagnostica und der Zeitschrift für Differentielle und Diagnostische Psychologie. (CDEP-Standards)

Die CDEP-Standards sind die deutsche Übersetzung der genannten amerikanischen Standards, die 1985 veröffentlicht wurden (es gab mehrere Vorgängerversionen, zurückdatierend bis ins Jahr 1954). Einführend wird die wachsende Relevanz von Tests insbesondere an Schulen, Hochschulen, Betrieben und Behörden herausgestellt. Wichtige an Tests Beteiligte werden aufgelistet (Entwickler, Anwender, Sponsoren, Leiter, Rezensenten und schließlich Probanden). Zweck der CDEP-Standards, ist es, „Kriterien für die Evaluation von Tests, Test-Praktiken und die Effekte der Testanwendung bereitzustellen.“

Die CDEP-Standards gelten in ihrer Gesamtheit, unterschieden wird nach Grad der (Un)Bedingtheit ihrer Anwendung:

- Primäre Standards, in jedem Falle anzuwenden
- Sekundäre Standards, wünschenswert aber oft zurückzustellen
- Situationsabhängige Standards, die je nach Kontext primär oder sekundär sein können.

Anzuwenden sind die CDEP-Standards auf Leistungstests, Einstellungsfragebogen, Inventare sowie „strukturierte Verhaltensproben“, darüber hinaus auf den gesamten Bereich der Erhebungsverfahren. Im Mittelpunkt stehen jedoch Tests im Sinne von Leistungsaufgaben sowie die Testdurchführung.

Es folgen vier Teile mit jeweils zwei bis sieben Kapiteln. Jedes der Kapitel enthält einen einführenden Text mit Hintergrundinformationen. Insgesamt gibt es 180 Standards, die aus einem oder mehreren Sätzen bestehen. Teilweise sind diese Standards kommentiert. Bei den Standards selbst handelt es sich um Aufforderungen, z.T. als Sollaussagen formuliert.

- Teil I umfaßt methodische Standards für die Test-Konstruktion und -Evaluation, in Kapiteln wie „Validität“, „Reliabilität und Meßfehler“ oder „Testveröffentlichung“.
- Teil II enthält die Standards für eine fachlich kompetente Testanwendung, insbesondere in unterschiedlichen Feldern wie klinisches Testen, Tests an Schulen, in der Beratung oder Berufseignungsdiagnostik. Hier ist auch ein Kapitel zur Programmevaluation enthalten.
- Teil III enthält Kapitel zum Testen bei sprachlichen Minderheiten und bei Personen mit Behinderungen.
- Teil IV enthält Standards für Vorgehensweisen, insbesondere „Test-Durchführung, -Auswertung und Ergebnisdarstellung“ sowie zum Schutz der Rechte der Probanden.

Es gibt außerdem ein acht Seiten umfassendes Glossar sowie einen aktuellen Aufsatz zum laufenden Revisionsprozeß der CDEP-Standards; mit einer Neuauflage des amerikanischen Originals ist im Jahr 2000 zu rechnen. Hervorzuheben sind die zahlreichen Bezüge der CDEP-Standards zur Programmevaluation. Das Kapitel 12 im Teil II enthält hierzu insgesamt acht Standards. Im Zusammenhang mit Evaluation wird grundsätzlich ein multimethodisches Vorgehen eingefordert – ein eventuell vorhandener Zusammenhang mit politischen und Finanzentscheidungen müsse sehr deutlich berücksichtigt werden. Es fällt überdies auf, daß mit George F. Madaus ein Evaluationstheoretiker Mitglied im Kernkomitee zur Erarbeitung der CDEP-Standards war – unter den Beratern finden sich so bekannte Evaluatoren wie Lee J. Cronbach, James Popham, Lee P. Sechrest und Daniel Stufflebeam. Schließlich werden die Standards des Joint Committee on Standards for Educational Evaluation von 1981 in der mit dreißig Titeln recht knappen Bibliographie aufgeführt.

Auffallend ist das Bemühen der CDEP-Standards um eine relativ hohe Allgemeinverständlichkeit: Auch Nicht-Testspezialisten sollen die formulierten Anforderungen nachvollziehen können. Mit 180 sind sechsmal so viele Standards formuliert wie in den ‚Standards‘. Die Tatsache, daß sehr viele als „primäre Standards“ ausgewiesen sind, verweist auf das Bemühen um eine vergleichsweise hohe Verbind-

lichkeit. Ein weiterer Unterschied besteht darin, daß im Teil II Standards für Tests in jeweils verschiedenen Anwendungsfeldern aufgeführt sind. Während ethische Standards in den Teilen III und IV größere Berücksichtigung finden, wird auf die Frage der Durchführbarkeit von Tests (Finanzeinsatz, ‚Störung‘ laufender Prozesse ...) nicht eingegangen.

Die CDEP-Standards sind eine wichtige Ergänzung der ‚Standards‘, in bezug auf den Einsatz von Tests als Datenerhebungsmethode in der Evaluation. Umgekehrt ergänzen die ‚Standards‘ die CDEP-Standards in bezug auf deren Einsatz im Kontext von Programmevaluationen.⁴⁶

2. Deutsche Forschungsgemeinschaft. Qualitätskriterien der Umfrageforschung. Denkschrift, 1999 (im Erscheinen) (DFG-Kriterien)⁴⁷

Gegenstand dieses insgesamt ca. 200 Seiten umfassenden Buches ist die Umfrageforschung, insbesondere die Wahl- und politische Meinungsforschung. Anlaß für diese Denkschrift sind (1) die durch die Möglichkeit von Telefoninterviews und die technische Realisierbarkeit von elektronischen Befragungen mittels des Internet veränderten technologischen Grundlagen der repräsentativen Umfrageforschung, (2) die mit den technologischen Möglichkeiten sinkenden Gründungskosten für Unternehmen, die bundesweit repräsentativ Umfragen durchführen können (schrumpfende Mindestbetriebsgrößen und damit verbundene Qualitätsprobleme), (3) das, wegen der erhöhten Taktgeschwindigkeit von Umfragen und ihrer massenmedialen Verbreitung, noch einmal verstärkte Interesse der Medien an solchen Ergebnissen und aufgrund der Rückwirkung auf die politische Meinungsbildung erhöhten Verantwortung des Gesamtsystems „Umfrageforschung und deren Verbreitung“.

Die Denkschrift richtet sich vorrangig an Umfrageforscher in Hochschuleinrichtungen und in privatwirtschaftlichen Umfrageinstituten, aber auch an die Massenmedien als Verbreitungsorgan sowie die universitären Ausbildungsstätten als für die Qualifikation des (künftigen) Forschungspersonals zuständig.

In dieser Denkschrift werden zahlreiche auch internationale Erfahrungen zusammengetragen, um Qualitätsstandards insbesondere für den Methodeneinsatz in der Umfrageforschung zu formulieren. Neben Begriffsdefinitionen finden sich insbesondere Qualitätskriterien zu Elementen wie Stichprobe, Offenlegung der Stichprobenrealisierung, Interviewerkontrollen. Auch Möglichkeiten prozeßorientierter Qualitätskontrollen (z.B. nach ISO 9000ff.) werden diskutiert.

Zum gegenwärtigen Zeitpunkt ist nicht absehbar, ob die Arbeitsgruppe eher Materialien zur Entwicklung von Qualitätskriterien im Bereich der Umfragefor-

46 Seit etwa 1995 läuft – von BDP und DGPs initiiert – ein Antrag beim Deutschen Institut für Normung, mit dem Ziel, etwa 2000 eine DIN-Norm für psychodiagnostische Instrumente zu verabschieden – eine Parallele zu den als ANSI-Normen anerkannten ‚Standards‘.

47 Zu dieser Denkschrift existieren zum Zeitpunkt des Redaktionsschlusses dieses Beitrages lediglich Entwürfe, die den Verfassern nur in Teilen zur Verfügung stehen.

schung bereitstellt oder ob zumindest ein Teil dieser Publikation auch konsensual verabschiedete Standards oder Anforderungen an Umfrageforschung und ihre Verbreitung darstellt.

Im Unterschied zu den ‚Standards‘ werden methodische Aspekte viel ausführlicher und differenzierter behandelt. Dies betrifft überwiegend quantitative, nach dem repräsentativen Zufallsstichprobenverfahren arbeitende Befragungen. Auch ethische Gesichtspunkte, insbesondere demokratietheoretische Fragen werden ausführlich behandelt (offensichtlich aber weniger solche, welche die Interaktion in der Befragungssituation betreffen). Fragen der Durchführbarkeit (z.B. Kostenaspekte) werden vergleichsweise kurz angesprochen. Aspekte der Vorbereitung von Nutzung, des Einbezugs von Beteiligten & Betroffenen fehlen – ggf. erklärbar durch den eindeutigen Schwerpunkt auf Umfrageforschung für Auftraggeber, die durch die exklusive Überlassung von Ergebnissen marktliche oder politische Vorteile erringen möchten. Umfrageforschung weist im Unterschied zur Evaluation keine Tradition in Programm-Kontexten auf, sondern agiert auf einem ökonomischen Markt von Produktanbietern und -nachfragern oder einem politischen Markt von Parteien und Wählern.

Es ist zu erwarten, daß sich aus dieser Denkschrift zahlreiche Anregungen ergeben, sowohl zur Ergänzung der Genauigkeits-‚Standards‘, insbesondere bezüglich quantitativer Methoden, als auch zur „Übersetzung“ der Korrektheitsstandards in deutschsprachige Demokratietraditionen.

3. Standards für die Weiterbildung zum/zur Supervisor/in, Deutsche Gesellschaft für Supervision e.V., Köln (1998) (DGSv-Standards)

Eine wichtige Funktion der Deutschen Gesellschaft für Supervision besteht darin, Ausbildungsgänge freier Institute (ca. dreißig sind dem Verband angeschlossen) zum Supervisor/zur Supervisorin anzuerkennen. Die Standards beschreiben Mindestanforderungen an solche Ausbildungen.

Als Eingangsvoraussetzungen sind an Teilnehmende solcher Ausbildungen zu stellen: ein einschlägiger Hochschulabschluß, minimal fünfjährige Berufserfahrung, dreißig Sitzungen Supervision sowie nachgewiesene Kenntnisse durch längerfristige Zusatzausbildungen – außerdem werden Ausnahmeregelungen aufgelistet.

Abschlußvoraussetzungen sind u.a. Teilnahme an Supervisions-Ausbildungskursen im Umfang von mind. 500 Unterrichtsstunden, drei eigene Lernsupervisionen à 15 Sitzungen, ca. fünfzig Stunden Teilnahme an Lehrsupervisionen sowie schriftliche und mündliche Leistungen.

An konzeptionellen Voraussetzungen wird die durchgängige Leitung der Supervisionsausbildungen durch DGSv-erkannte Supervisoren/-innen gefordert, das Vorhandensein einer schriftlichen Ausbildungskonzeption mit persönlichkeits-theoretischen und didaktischen Festlegungen, sowie bestimmte Anforderungen an Mindestdauer und Elemente solcher Ausbildungskurse.

Die DGSv-Standards regulieren insbesondere die Zulassung von Ausbildungskonzepten, die im Rahmen der DGSv anerkannt werden können, die Mindesteingangsvoraussetzungen für Teilnehmende, den Mindestumfang der Teilnahme. Bei gleichzeitig gegebener möglicher Vielfalt persönlichkeits-theoretischer Zugänge wird eine formale Standardisierung der Ausbildungsprozesse mit gezielten Eingangshürden (Hochschulabschluß sowie Erfahrung als Supervisand bei verbands- anerkannten Supervisoren/innen) geschaffen.

Im Vergleich zu den ‚Standards‘ handelt es sich um typische Minimalstandards, die klar operationalisiert sind und deren Einhaltung fortlaufend überwacht wird. Die Standardisierung zielt auf die Kompetenzen der Ausgebildeten; die Steuerung erfolgt über die Zulassung von Ausbildungsinstituten. Die Sanktionsmöglichkeiten des Verbandes sind vergleichsweise hoch, insofern Nicht-Einhaltung der Ausbildungsrichtlinien zur Annullierung der Anerkennung führen kann, was einen erheblichen Marktnachteil für die Institute darstellt.

4. Empfehlungen zur Supervision – Bedeutung und Organisation. Herausgegeben vom Deutschen Verein für öffentliche und private Fürsorge in seinem Nachrichtendienst, 73. Jahrgang 1993, Heft 3, S. 89-92 (dv-Empfehlungen).

Die Empfehlungen wurden durch den Vorstand des Deutschen Vereins für öffentliche und private Fürsorge (dv) im Dezember 1992 verabschiedet. Der dv ist der zentrale Zusammenschluß der öffentlichen und freien Träger der sozialen Arbeit in der Bundesrepublik Deutschland. Die dv-Empfehlungen wenden sich in erster Linie an die Anstellungsträger von Fachkräften in der sozialen Arbeit. Für erforderlich werden diese Empfehlungen gehalten, da sich soziale Arbeit „weitgehend den klassischen Instrumenten einer Wirksamkeitsprüfung entzieht“ und damit „Supervision *ein* geeignetes und bewährtes Instrument der Steuerung und mittelbar der Überprüfung der fachlichen Arbeit“ ist. Mit diesen Empfehlungen wird Supervision in gewisser Weise als zur Evaluation alternatives Qualitäts-Steuerungsverfahren für die soziale Arbeit beschrieben.

Supervision wird definiert als das berufliche Handeln begleitender, begrenzter, kontinuierlicher Lehr- und Lernprozeß. Unterschieden werden interne und externe Supervision.

Gefordert wird, zur Inanspruchnahme von Supervision als „fester Bestandteil sozialer Arbeit“ zu motivieren. Damit dies gelingt, ist hoher Vertrauensschutz zu gewährleisten. Außerdem soll zwischen Träger, Supervisanden und Supervisor eine Vereinbarung zur Supervision getroffen werden, für die wünschbare Bestandteile aufgeführt werden. Qualität von Supervision wird gewährleistet (1) durch verantwortliches Leitungshandeln des Auftraggebers, (2) durch qualifiziert ausgebildete Supervisoren/Supervisorinnen mit Verweis auf die DGSv-Standards (siehe oben).

Im Unterschied zu den ‚Standards‘ liegt der Schwerpunkt auf Definitionen verschiedener Arten des Leistungsprozesses (Supervision). Auch die Forderung nach einer zertifizierten Supervisionsausbildung und der Verzicht auf Schriftlichkeit von

Auswertungen markieren Differenzen. Überschneidungen gibt es beim Schließen von Vereinbarungen oder der Vertraulichkeit.

5. Government Auditing Standards, 1994 Revision. United States General Accounting Office, Comptroller General of the United States. June. Washington: GAO, June 1994 (GA-Standards)

Die ‚Government Auditing Standards‘ (GA-Standards) sind vom ‚General Accounting Office‘ (GAO), der Kontroll- und Aufsichtsbehörde des US-Kongresses herausgegeben und auch unter der Bezeichnung ‚the yellow book‘ bekannt. Das gelbe Buch umfaßt in der Fassung aus dem Jahre 1994 insgesamt 111 Seiten. Es enthält Vorschriften zur Durchführung von zwei Typen von Überprüfungen („audits“): von Finanzkontrollen („financial audits“) und von Leistungskontrollen („performance audits“). Erstere orientieren sich an den Regeln der Rechnungsprüfung bzw. Rechnungslegung und beschränken sich auf finanzielle Aspekte. Die Normen zu den Leistungsüberprüfungen entstammen zwar auch der Kultur der Rechnungsprüfung, liegen aber in ihrem Gehalt der Evaluation deutlich näher (vgl. dazu Davis 1990). Die Bindung an die Kultur des Rechnungswesens zeigt sich etwa daran, daß Kriterien der Wirtschaftlichkeit, Effizienz und Ordnungsmäßigkeit deutlich stärker vertreten sind, als dies bei den ‚Standards‘ der Fall ist.

Die GA-Standards umfassen sowohl allgemeine Standards, die sich auf Finanz- und Leistungsüberprüfungen beziehen als auch spezifische Normen für die beiden Untersuchungstypen. Allgemeine Normen sind u.a.: Qualifikation und Unabhängigkeit des eingesetzten Personals, gebührende professionelle Sorgfalt und Qualitätskontrolle. Die für die Leistungsüberprüfungen spezifischen Normen sind unterteilt in solche, die sich mit der Durchführung und solche, die sich mit der Berichterstattung befassen. Für die Durchführung lauten sie: Adäquate Planung, Aufsicht, Einhalten der Gesetze und Vorschriften, Managementkontrolle sowie ausreichende, kompetente und relevante Evidenz. Die Normen zur Berichterstattung heißen: Form (Schriftlichkeit), Rechtzeitigkeit, Vorgaben zum Berichtsinhalt, Präsentation des Berichts, Verbreitung des Berichts.

Vergleicht man die GA-Standards mit den ‚Standards‘ ist in erster Linie darauf hinzuweisen, daß die GA-Standards nicht nur Leistungsüberprüfungen regeln, sondern auch Finanzprüfungen. Beschränken wir den Vergleich auf den Bereich der Leistungskontrolle, sind einige Parallelen, aber auch deutliche Differenzen feststellbar. Die Vorgaben zur Aufsicht und zur Qualitätskontrolle erinnern an den Standard zur Meta-Evaluation. Sehr ähnlich sind die allgemeinen Vorgaben zur Rechtzeitigkeit oder zur Berichterstattung. Betrachtet man die Vorgaben der GA-Standards zur Berichterstattung genauer, fällt auf, daß sie deutlich präskriptiver sind. So bestimmen die GA-Standards detailliert, welche Themen im Bericht enthalten sein müssen. Einen deutlichen Kontrapunkt zur Orientierung der ‚Standards‘ an den Beteiligten & Betroffenen bildet das Gebot der Neutralität/Unabhängigkeit der GA-Standards. Die ‚Standards‘ fordern zwar von der Evaluatorin eine gewisse Distanz zum Evaluationsgegenstand, aber viel weniger strikt als die GA-Standards.

Bezüglich der Untersuchungsmethodik sind die ‚Standards‘ deutlich detaillierter und zudem methodologisch offener als die GA-Standards. Eine weitere wichtige Differenz betrifft die Geltungskraft der beiden Regelwerke: Die GA-Standards stellen verbindliche Vorschriften der obersten Kontrollinstanz der USA dar (Comptroller General of the United States). Dadurch kommt ihnen eine deutlich höhere Verbindlichkeit zu als den ‚Standards‘.

*6. Berufsethische Normen. Vereinigung Schweizer Markt- und Meinungsforschungsinstitute (Swiss Interview)/Verband Schweizer Marketing- und Sozialforscher (SMS), o.J., (SMS-Normen)
und
Richtlinien zur Durchführung von abstimmungs- und wahlbezogenen Umfragen, die zur Veröffentlichung vor dem Urnengang bestimmt sind. Vereinigung Schweizer Markt- und Meinungsforschungsinstitute (Swiss Interview), 12. Mai 1993. (Swiss Interview-Richtlinien)*

Die berufsethischen Normen des SMS und der Swiss Interview (SMS-Normen) sowie die Richtlinien der Swiss Interview für spezifische Umfragen (Swiss Interview-Richtlinien) sind in zwei separaten Texten festgehalten.

Bei den *SMS-Normen* handelt es sich um ein Merkblatt im Umfang von zwei Seiten. Es werden folgende Aspekte behandelt: Pflichten des Forschers gegenüber den Auftraggebern, Forschungsvorschläge und Offerten, Berichterstattung und Präsentation von Untersuchungsergebnissen, Pflichten des Forschers gegenüber den Befragten, Allgemeine Bestimmungen für Offertstellung. Die Verbandsmitglieder sind durch diese Normen gebunden und haben sich dazu verpflichtet, sie zu befolgen. Die Normen befassen sich vor allem mit dem Verhältnis des Forschers zu Auftraggeber und Befragten. Dabei handelt es sich primär um Vorgaben, die ein redliches Geschäftsgebaren oder ethische Fragen betreffen. Ein drittes Hauptthema bildet die Berichterstattung, wobei hier detaillierte Angaben über die einzuhaltenen, minimalen Anforderungen enthalten sind. Besonderes Augenmerk gilt – im Vergleich zu anderen, hier präsentierten Regelwerken wie auch zu den ‚Standards‘ – der Regelung von Publikationsrechten an Ergebnissen. Die SMS-Normen übertragen dem Auftraggeber das alleinige Recht, über eine Publikation der Resultate zu entscheiden. Der Auftragnehmer hat demnach ohne explizite Einwilligung des Auftraggebers auf eine Publikation der Ergebnisse zu verzichten.

Die *Swiss Interview-Richtlinien* regeln einen spezifischen Bereich der Aktivitäten der Verbandsmitglieder. Sie beziehen sich ausschließlich auf allgemeine Bevölkerungsbefragungen zu Abstimmungen und Wahlen, die zudem vor dem Urnengang veröffentlicht werden sollen. Die Richtlinien bilden einen festen Bestandteil für Offerten, Verträge und Berichte der Mitglieder. Die Richtlinien enthalten verbindliche Angaben zu den folgenden Themen: Stichprobengröße, Art der Umfrage, obligatorische Informationen (an den Auftraggeber), Selbstbeschränkung bei der Durchführung (terminliche Vorgaben in Relation zum Urnengang).

Sowohl die berufsethischen SMS-Normen wie auch die Swiss Interview-Richtlinien werden zur Zeit überarbeitet. Die Organisation Swiss Interview existiert heute nicht mehr selbständig, sondern ist Teil der SMS. Swiss Interview soll hingegen künftig eine sog. Kollektivmarke der SMS-Institute darstellen. Mit der Signalisierung als geschützter Markenartikel nach außen wird die Verbindlichkeit der Normen für die Mitglieder erhöht. Bei den Arbeiten an diesen neuen Grundlagen stützt sich die SMS auf Normen der European Society for Opinion and Marketing Research (ESOMAR)⁴⁸ sowie auf Unterlagen des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM)⁴⁹.

Die Swiss Interview-Richtlinien haben, im Vergleich zu den ‚Standards‘, einen engen, sehr spezifischen Geltungsbereich. Sie beziehen sich ausschließlich auf standardisierte Befragungen (und nicht auf andere Methoden der Datenerhebung) zu politischen Themen, über die eine Volksabstimmung erfolgen wird, und zwar nur auf solche Umfragen, bei denen auch eine Publikation der Ergebnisse vor dem entsprechenden Urnengang erfolgen soll. Diese Einschränkungen haben nicht nur zur Folge, daß der Geltungsbereich eingeschränkt wird, sondern auch, daß äußerst präzise Angaben für ein korrektes Verhalten der Verbandsmitglieder möglich sind.

III. Evaluationsunspezifische Regelwerke, die auf Fachpersonen gerichtet sind

1. Ethical Standards of the American Educational Research Association. American Educational Research Association, June 1992 (AERA-Standards)

Die Ethik-Standards der ‚American Educational Research Association‘ (AERA-Standards) haben einen Umfang von sieben Seiten. Darin werden Leitlinien in sechs Themengruppen aufgeführt: 1. Verantwortung gegenüber dem Forschungsfeld; 2. Untersuchungsgruppen, Bildungsinstitutionen und Öffentlichkeit; 3. Urheberrechte; 4. Herausgabe, Begutachtung und Beurteilung von Forschung; 5. Sponsoren, politische Entscheidungsträger und andere Nutzer der Forschungsergebnisse; 6. Studierende und Jungforscher. Die AERA-Standards wollen eine Selbstreflexion der Mitglieder anregen, kollegiale Debatten stimulieren und eine freiwillige Einhaltung durch Überzeugung erreichen. Die AERA-Standards sind sehr vielfältig. Die sechs Themengruppen enthalten jeweils eine größere Zahl von kurzgefaßten Einzelstandards.

Im Gegensatz zu disziplinär ausgerichteten Regelwerken sind die Standards der AERA disziplinenübergreifend angelegt, der breiten Rekrutierung ihrer Mitglied-

48 Siehe <http://www.esomar.nl/>, bes. die ESOMAR Guidelines.

49 Siehe dazu <http://www.adm-ev.de/>. Dabei ist im Zusammenhang mit den ‚Standards‘ besonders auf die „Standards zur Qualitätssicherung in der Markt- und Sozialforschung“ des ADM hinzuweisen.

schaft entsprechend. Die Gemeinsamkeit bildet der Untersuchungsgegenstand „Bildung und Erziehung“.

Inhaltlich fällt bei den AERA-Standards das große Gewicht auf, das den Beziehungen zu Personen und Institutionen im Untersuchungsfeld zugemessen wird. Eine weitere Spezialität sind die Themen Herausgabe, Begutachtung und Beurteilung von Forschungsergebnissen, die sich vornehmlich auf AERA-eigene Publikationen beziehen. Die Adressaten der AERA-Standards sind primär die Mitglieder der AERA, wobei einzelne Vorgaben auch beispielsweise die Zeitschriften der AERA binden. Die AERA hat keine Vorkehrungen zur Überwachung getroffen und nimmt auch keine Abklärungen vor, wenn ein Verdacht auf einen Verstoß gegen ihre Standards besteht.

Insgesamt markiert die Ausrichtung auf die Mitglieder bzw. die Publikationen der AERA einen deutlichen Unterschied zu den ‚Standards‘, bei denen eine solche Basis nicht (oder nur beschränkt – vermittelt über die unterstützenden Organisationen) besteht. Ebenso von Bedeutung ist der unterschiedliche Geltungsbereich der beiden Regelwerke. Die AERA-Standards beschränken sich auf ethische Fragen, während die ‚Standards‘ auch andere Themengebiete, wie etwa die Methodik abdecken.

2. A Guide to Professional Ethics in Political Science. 2nd Edition. American Political Science Association, Committee on Professional Ethics, 1991 (APSA-Guide).

Der APSA-Guide stellt den ‚Ethik-Führer‘ der amerikanischen politikwissenschaftlichen Vereinigung dar, also einer disziplinären Fachvereinigung mit ausgeprägter akademischer Orientierung. Er ist 1991⁵⁰ in zweiter Ausgabe erschienen (Publikation der Erstfassung: 1968). Der Guide umfaßt insgesamt 35 Seiten und bildet die Arbeitsgrundlage für eine durch die Organisation eingesetzte Kommission, die sich mit Verstößen gegen die ethischen Prinzipien befaßt.

Da die Mitglieder der APSA überwiegend an Hochschulen arbeiten, enthält der Guide umfangreiche Vorgaben zum universitären Tätigkeitsfeld. Integraler Bestandteil des Guide sind die Bestimmungen zur professionellen Ethik der American Association of University Professors. Inhaltliche Schwerpunkte liegen bei Themen aus dem Universitätsbetrieb wie Lehre, Personalfragen, Publikationen und Forschung, wobei den Forschungstätigkeiten im engeren Sinne ein verhältnismäßig kleiner Teil gewidmet ist.

Die Konzentration auf die Hochschule und die geringe Berücksichtigung von Forschungs- oder Untersuchungsaktivitäten markieren deutliche Unterschiede zu den ‚Standards‘. Die Normen des Guide richten sich an den (universitären) Politikwissenschaftler als Person (oder an politikwissenschaftliche Abteilungen als Orga-

50 Der APSA-Guide ist seither mehrfach in einzelnen Punkten revidiert worden – eine Broschüre mit der aktuellen Fassung soll demnächst erscheinen. Der APSA-Guide findet sich in aktueller Fassung im WWW unter der Adresse <http://www.apsanet.org/PS/ethics.html>.

nisationen). Damit umschreiben sie ein ethisch wünschbares Verhalten dieser Akteure, nicht die ethische Qualität der Politikwissenschaft als Fach oder Disziplin bzw. politikwissenschaftlicher Studien. Schließlich fokussiert der Guide ethische Fragen, während diese bei den ‚Standards‘ lediglich einen kleinen Teil ausmachen (insbesondere Standardgruppe „Korrektheit“). Der Guide enthält deutlich spezifischere Angaben zu ethischen Anforderungen. Ein weiterer Unterschied liegt darin, daß der APSA-Guide nicht nur Prozeduren zu dessen Weiterentwicklung vorsieht, sondern auch zu dessen Durchsetzung bei den Angehörigen der Fachdisziplin.

3. Verhaltenskodex für zertifizierte Unternehmensberater „CMC“. Schweizerische Vereinigung der Unternehmensberater (ASCO), Zertifizierungskommission „CMC“ (Certified Management Consultant). Zürich, September 1996. (CMC-Kodex)

Der CMC-Kodex wurde von der Zertifizierungskommission der Schweizerischen Vereinigung für Unternehmensberater (Association Suisse des conseils en organisation et gestion, ASCO) erarbeitet. Der Inhalt des CMC-Kodex bildet einen Bestandteil der Zertifizierungsüberprüfung. Verstöße gegen den Kodex können dazu führen, daß einem Berater das Zertifikat entzogen wird. Der Kodex soll regelmäßig überarbeitet resp. ergänzt werden.

Der CMC-Kodex enthält insgesamt zehn Abschnitte zu den folgenden Themen: Vertraulichkeit, unrealistische Erwartungen, Kommissionen/finanzielle Interessen, Aufträge, unverträgliche Aufträge, Verhandlungen mit dem Klienten, Rekrutierung, Vorgehen, andere Unternehmensberater und zusätzliche Interpretationen. Diese Abschnitte enthalten jeweils einen bis fünf Unterabschnitte (Interpretationen genannt), in denen der Gehalt der verschiedenen Normen verdeutlicht und spezifiziert wird. Insgesamt umfaßt der CMC-Kodex vier Seiten Text.

Im Vergleich des CMC-Kodex mit den ‚Standards‘, fällt vor allem auf, daß dieser

- Bestandteil eines Zertifizierungsprozesses ist;
- großes Gewicht auf die Beziehung zum Auftraggeber sowie auf potentielle Interessenkonflikte legt;
- auf Normen zur Methodik der Vorgehensweise verzichtet und sich auf ethische Fragen beschränkt.

4. Berufsethische Prinzipien des Deutschen Berufsverbandes für Sozialarbeit, Sozialpädagogik und Heilpädagogik e.V., Essen 1997 (DBSH- Prinzipien)

Die knapp zwanzig Seiten umfassende Broschüre des Deutschen Berufsverbandes für Sozialarbeit, Sozialpädagogik und Heilpädagogik e.V. (DBSH; ca. 6.000 Mitglieder) enthält neben den eigenen, im November 1997 beschlossenen DBSH-Prinzipien den Abdruck zweier weiterer Normenwerke für Sozialarbeiter/innen, So-

zialpädagogen/-innen und Heilpädagogen/-innen. Zum einen ist dies ein Ausschnitt aus dem „Internationalen Ethikkodex für die beruflichen SozialarbeiterInnen“, der bereits 1976 von der International Federation of Social Workers verabschiedet wurde (IFSW-Standards). Das dritte Dokument gibt einen 1994 aktualisierten Abschnitt aus den älteren IFSW-Standards wieder: Die „Ethischen Grundlagen der sozialen Arbeit – Prinzipien und Standards“ (IFSW-Prinzipien).

Das Besondere der DBSH-Prinzipien liegt darin, daß sie explizit im Zusammenhang mit den Prinzipien einer internationalen Berufsvereinigung erarbeitet worden sind. So verpflichten die IFSW-Prinzipien die einzelnen nationalen Mitgliedsverbände darauf, in ihrem jeweiligen Zuständigkeitsbereich Ethikforen für Streitfragen und Probleme einzurichten und eigene ethische Standards insbesondere für Arbeitsbereiche zu entwickeln, bei denen komplizierte ethische Probleme zu erwarten sind.

Die DBSH-Prinzipien formulieren unter der Überschrift „Allgemeine Grundsätze beruflichen Handelns“ wertgebundene Positionen, die Sozialarbeiter einnehmen und für die sie sich einsetzen sollen: Dies gilt beispielsweise für das Prinzip des freien Zugangs zu allen Sozialleistungen ohne Diskriminierung, die Bekämpfung strukturell bedingter Ursachen sozialer Not oder das Eintreten für die Rechte sozial Benachteiligter. Die nachfolgenden Hauptabschnitte richten sich – in ‚konzentrischen Kreisen‘ aufgebaut – auf das Verhalten (a) gegenüber der eigenen Klientel, (b) gegenüber BerufskollegInnen, (c) gegenüber Angehörigen anderer Berufe, (d) gegenüber ArbeitgeberInnen und schließlich (e) in der Öffentlichkeit.

Bezüglich der Klientel wird insbesondere darauf hingewiesen, daß personenbezogene Daten nur soweit nötig gespeichert werden dürfen, daß mit Ihnen strikt vertraulich umzugehen ist und daß Klienten möglichst weitgehend Zugang zu den Unterlagen haben sollen (deutliche Unterschiede etwa gegenüber den DGBS/BDP-Richtlinien).

Es wird von den Mitgliedern gefordert, daß sie die in Ausübung ihres Berufes gewonnenen Erkenntnisse und getroffenen Maßnahmen dokumentieren und zur Planung und Reflexion des Arbeitsprozesses einsetzen. Bezüglich eines auch für Außenstehende nachvollziehbaren Verfahrens der Qualitätsentwicklung bleiben die DBSH-Prinzipien damit hinter den IFSW-Standards zurück, welche fordern: „Legen Sie Rechenschaft ab gegenüber KlientInnen und der Gesellschaft in Form von Tätigkeitsberichten, welche die Wirtschaftlichkeit und Wirksamkeit der professionellen Dienstleistung aufzeigt.“

Kennzeichnend für alle drei Regelwerke ist die starke Wertorientierung, die Positionsnahme für sozial Schwache und Benachteiligte und das Eintreten für soziale Gerechtigkeit.

5. Ethische Richtlinien der Deutschen Gesellschaft für Psychologie e.V. und des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V. 1998 (DGPs/BDP-Richtlinien)

Dieses Papier ist aus einer gemeinsamen Initiative der akademisch orientierten Deutschen Gesellschaft für Psychologie und des berufsständische Interessen vertretenden Berufsverbandes hervorgegangen. Während die Mitgliederversammlung der DGPs sie bereits im September 1998 verabschiedet hat, stand zum Redaktions-schluß die Bestätigung durch die Delegiertenkonferenz des BDP noch aus. Von besonderer Bedeutung ist, daß diese Ethikrichtlinien zugleich Berufsordnung des BDP sind.

In der Präambel wird als primäres Ziel der Richtlinien genannt: „Der Schutz und das Wohl der Menschen, mit denen Psychologen arbeiten.“ Es werden breite Geltungsbereiche aufgeführt, angefangen bei Lehre und Forschung, über Diagnostik, Psychotherapie, Beratung, Supervision u.a.

Kennzeichnend für diese Ethikrichtlinien ist eine doppelte Orientierung: Einerseits an dem neuesten Stand der wissenschaftlichen Erkenntnisse, an denen sich alle beruflichen Handlungen von Psychologen ausrichten sollen, andererseits an einem Berufsleitbild für Psychologen, das dem ‚freien Beruf‘ verpflichtet ist und die Nähe zum Berufsbild des Arztes sucht. So findet sich in den allgemeinen Bestimmungen die Verpflichtung auf ein abgeschlossenes Hauptfachstudium in Psychologie als Eingangsvoraussetzung in den Berufsstand. Entsprechend der hohen personalen und professionellen Anforderungen werden Psychologen auf fortwährende Fortbildung verpflichtet.

In seinen zahlreichen Abschnitten und Unterabschnitten sind oft sehr genaue Soll-Aussagen formuliert, die ein erwünschtes ethisches Verhalten von Psychologen⁵¹ bisweilen hochgradig operationalisiert beschreiben. So werden z.B. präzise Aussagen zur Schweigepflicht, zum Umgang mit Daten, zu den Grundsätzen der Werbung und dem erwünschten Verhalten von Psychologen bei Auftritten in der Öffentlichkeit gemacht.

Im Hauptabschnitt „Psychologie und Forschung/Lehre“ wird das Idealbild eines von Fremdbestimmung und wissenschaftsfremder Parteilichkeit frei handelnden Psychologen gezeichnet, u.a. mit Verweis auf das Grundrecht der Wissenschaftsfreiheit (Artikel 5 des Grundgesetzes). Das Stichwort „Redlichkeit“ erinnert an die Gruppe der Korrektheits-„Standards“. So wird z.B. gefordert, daß Forschungsergebnisse nicht nur dokumentiert, sondern umfassend zugänglich gemacht werden, daß auch den eigenen Auffassungen widersprechende Daten veröffentlicht werden und Korrekturen eingeleitet werden, wenn nachträglich bedeutsame Fehler in veröffentlichten Werken festgestellt werden.

Der Hauptabschnitt „Psychologie in der Anwendung“ gilt zwar für klinisch tätige, insbesondere heilkundlich tätige Psychologen. Stichworte wie „Vertrauensverhältnis“, „Aufklärung und Einwilligung“, „Wahrung der Unabhängigkeit“ zeigen jedoch an, daß diese Anforderungen auch für in der Forschung tätige Psychologen Relevanz haben.

51 Auf ein wichtiges europäisches Regelwerk wird in diesem Artikel nicht Bezug genommen, den „European Meta-Code of Ethics and Carta Ethica“ der European Federation of Professional Psychologists Associations, 1995 in Athen durch die Generalversammlung der EFPPA verabschiedet.

Die Ethikrichtlinien sind in einer Spannung zwischen ausgeprägter Wissenschaftsorientierung einerseits und den zur Durchsetzung berufsständischer Interessen notwendigen Regulierungen andererseits entstanden. Sie zeichnen beispielhaft vor, welche Festlegungen getroffen werden müssen, wenn eine Profession über operationale Ein- und Ausschlusskriterien von Mitgliedern konstituiert wird. Eventuelle Überlegungen, im Bereich von „Evaluation“ berufsständische Bestimmungen oder Zertifizierungen einzuführen, sollten sich mit diesen Ethikrichtlinien intensiv auseinandersetzen. Zahlreiche Differenzierungen finden sich zu Aspekten, die in den ‚Standards‘ unter den Stichworten „Schutz individueller Menschenrechte“ (K3) sowie „Human gestaltete Interaktionen“ (K4) angesprochen sind.

6. *Ethik-Kodex der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologen (BDS)*. In: *Sozialwissenschaften und Berufspraxis*, Jg. 16, Nr. 1 (Januar/März 1993): 61-66, (DGS/BDS-Kodex)
Ethik-Kodex der Deutschen Vereinigung für Politische Wissenschaft. DVPW, o.J. (DVPW-Kodex) und
Ethik-Kodex der Schweizerischen Vereinigung für Politische Wissenschaft. Schweizerische Vereinigung für Politische Wissenschaft (SVPW), 1995. In: SVPW-Bulletin 90, April 1996. (SVPW-Kodex)

Die drei genannten Ethik-Kodizes sind in der angeführten Reihenfolge entstanden und stützen sich stark aufeinander ab. Der zuerst geschaffene Kodex der deutschen soziologischen Profession ist in einer gemeinsamen Arbeitsgruppe der Deutschen Gesellschaft für Soziologie (DGS), des Berufsverbandes Deutscher Soziologen (BDS) und der Gesellschaft für Soziologie (GfS) entstanden. Die GfS ist Ende 1992 aufgelöst worden – es handelte sich bei ihr um die Vereinigung der Soziologen aus der DDR. Der Kodex umfaßt einen Text von sechs Seiten mit den folgenden Abschnitten: Präambel; Forschung; Publikationen; Begutachtung; Der berufliche Umgang mit Studierenden; Mitarbeiter/innen und Kollegen/innen; Die Ethik-Kommission. Wie diese Inhaltsübersicht zeigt, ist dieser Ethik-Kodex im Vergleich zum vorher diskutierten APSA-Guide weniger stark auf das universitäre Leben ausgerichtet. Er enthält zwar einige Normen, die sich insbesondere auf die Lehrtätigkeit beziehen; im Vergleich zum umfangreichen Abschnitt zum Forschungsbereich kommt diesen aber ein kleineres Gewicht zu. Der DGS/BDS-Kodex sieht die Schaffung einer Ethik-Kommission vor, die einerseits die Aufgabe hat, zur Einhaltung der im Ethik-Kodex enthaltenen Vorgaben beizutragen und andererseits, gegebenenfalls Vorschläge zur Weiterentwicklung der Normen zuhanden der Verbände zu formulieren.

Der Ethik-Kodex der Deutschen Vereinigung für Politische Wissenschaft (DVPW-Kodex) lehnt sich stark an den DGS/BDS-Kodex an. Ein Vergleich der beiden Papiere ergibt geringe Differenzen. So verzichtet der DVPW-Kodex auf die im DGS/BDS-Kodex enthaltene Forderung, daß die Forschungsergebnisse wenn immer möglich veröffentlicht werden sollen.

Der Ethik-Kodex der Schweizerischen Vereinigung für Politische Wissenschaft (SVPW-Kodex) stützt sich seinerseits auf den DVPW-Kodex ab und übernimmt dessen Grobstruktur. Die vorgenommenen Änderungen gehen etwas weiter als bei der DVPW. Es sind insbesondere folgende Anpassungen vorgenommen worden:

- Generell formuliert der SVPW-Kodex knapper als jener der DVPW; mehrere Normen des DVPW wurden verkürzt übernommen.
- Die bei der DVPW durch eine Ethik-Kommission wahrgenommenen Aufgaben werden im SVPW-Kodex dem Vorstand dieser Vereinigung übertragen.
- Der Begriff der ‚Objektivität‘, der sowohl im DGS/BDS-Kodex wie auch beim DVPW-Kodex Verwendung findet, wird im SVPW-Kodex durch den Begriff ‚Sachlichkeit‘ ersetzt.

Vergleicht man diese drei Kodizes mit den ‚Standards‘, sind insbesondere folgende Unterschiede festzustellen:

- Die drei beschriebenen Ethik-Kodizes befassen sich ausschließlich mit ethischen Fragen, im Gegensatz zu den ‚Standards‘, die auch andere Themenfelder ansprechen.
- Die Ethik-Kodizes nehmen explizit Bezug auf die universitäre Lehrtätigkeit, wenn auch nicht so stark, wie dies beim APSA-Guide der Fall ist. Die ‚Standards‘ befassen sich dagegen nicht mit Fragen der Lehre (als Tätigkeit von Evaluationsfachleuten).
- Die Ethik-Kodizes enthalten im Gegensatz zu den ‚Standards‘ Angaben zur Durchsetzung der Normen.

7. Berufsordnung der Deutschen Gesellschaft für Supervision e.V., Köln (1997) (DGSv-Berufsordnung)

Die Berufsordnung der Deutschen Gesellschaft für Supervision (DGSv) umfaßt eine Präambel sowie zehn recht kurz gehaltene Paragraphen. Einleitend wird der Gegenstand von Supervision umrissen. Die Mitglieder der DGSv sind aufgefordert, sozial- und rechtsstaatlich verantwortlich zu handeln und für die Menschenwürde einzutreten. Sie sollen emanzipatorische Entwicklungen unterstützen und aufklärend insbesondere in bezug auf Macht und Konflikte im Bereich von Arbeitsbeziehungen wirken. Die ca. 2500 Mitglieder der DGSv werden auf ein Handeln im Rahmen der professionellen Standards und Einhaltung von Rechtsvorschriften verpflichtet. Als Mitglied dürfen sie die Bezeichnung Supervisor/-in DGSv führen. Kontraktschluß und Vertrauensschutz sind ebenso zu sichern wie kollegialer Umgang. Es werden minimale Regelungen bezüglich Fortbildung und regelmäßiger fachlicher Kontrolle sowie Werbung genannt. Bei angezeigten Verstößen gegen die Berufsordnung benennt der Verbandsvorstand einen Konfliktmoderator, der einvernehmliche Lösungen vorbereitet. Auffallend an diesem Regelwerk ist die Kürze trotz vorgesehenem Konfliktregelungsmechanismus, der sehr unbürokratisch konzipiert ist.

8. Ethische Rahmenrichtlinien der DGVT und deren Kommentare. Deutsche Gesellschaft für Verhaltenstherapie, 1996 (DGVT-Richtlinien)

Die DGVT-Rahmenrichtlinien schaffen Orientierung für die ca. 5000 Mitglieder dieses in psychosozialen Arbeitsfeldern angesiedelten Verbandes. Sie bestehen aus insgesamt sieben kurzen Absätzen, zu denen jeweils Kommentare bis zum Umfang einer Textseite gegeben werden.

Die Präambel weist aus, daß die Rahmenrichtlinien vorrangig zu einem Selbstvergewisserungsprozeß über gute psychosoziale Praxis beitragen sollen, wobei sie gleichzeitig als veränderlich und auf westeuropäische Kultur und Denkweise begrenzt angesehen werden.

Prozeßorientierung, ständiges Lernen und kritische Haltung gegenüber sich selbst kennzeichnen die Rahmenrichtlinien, erkennbar an Stichworten wie „Offenheit in der Auseinandersetzung und Begegnung“, „Bewußtheit über die Grenzen der Wahrnehmung“, „ständige Reflexion von Rolle, Auftrag oder Gestaltung in Wechselbeziehungen“. Gefordert wird Klarheit und Transparenz im professionellen Kontakt. Dem Dialog der Lebenswelten und Kulturen ist einer der Absätze gewidmet.

Die DGVT-Rahmenrichtlinien sind insgesamt ein visionäres, auf ständige Entwicklung angelegtes Leitbild, das deutlich wertgebundene Positionsbestimmungen enthält (z.B. die Betonung interkultureller, geschlechtspezifischer Perspektiven, die Forderung nach einer „Fähigkeit der Parteilichkeit ohne Schuldzuweisung“ oder die Norm, wann immer möglich Mitwirkungsbereitschaft und Einverständnis der Klienten einzufordern).

Der „Qualitätssicherung“ ist ein eigener Absatz gewidmet. Als Instrumente werden insbesondere genannt: Fortbildungsmöglichkeiten, Supervision und interdisziplinäre Zusammenarbeit. Folgender Satz im Kommentar bietet eine Brücke zur Evaluation, die in anderen deutschsprachigen Ethikrichtlinien so deutlich nicht zu finden ist: „Es ist notwendig, Kriterien zu entwickeln, die die Qualität der Arbeit prüfen. Solche Kriterien könnten sein: Zufriedenheit der KundInnen, Wirtschaftlichkeit, regelmäßige Überprüfung vereinbarter Ziele und der Grad ihrer Verwirklichung.“

IV. Evaluationsspezifische Regelwerke, die auf Leistungsprozesse gerichtet sind

1. Normen van de Algemene Rekenkamer voor de beoordeling van rapporten. Den Haag: Algemene Rekenkamer, 6 juni 1996: 5-6. (AR-Normen)

Die Normen der Algemene Rekenkamer, dem unabhängigen niederländischen Rechnungshof, sind im Zusammenhang mit einem konkreten Überprüfungsprojekt entstanden. Im Unterschied zu den anderen hier vorgestellten Normen stellen sie keine selbständige Publikation dar. In dem erwähnten Projekt (Algemene Rekenkamer 1996) ging es darum, die Qualität von insgesamt 49 Evaluationen zu überprüfen, die im holländischen Wirtschaftsministerium entstanden sind. Die Meta-Evaluation kommt zum Ergebnis, daß die technische Qualität der Berichte deutlich weniger überzeugt als ihre Nützlichkeit (vgl. dazu auch Widmer 1996a: 305). Die Kriterien sind in die zwei Gruppen der technischen Qualität und der Brauchbarkeit unterteilt. In der ersten Gruppe enthalten sind folgende Kriterien: Aktualität und Relevanz, Auswahl der Untersuchungseinheiten (Stichprobenziehung, Falldefinition, Repräsentativität etc.), Validität und Reliabilität der Instrumente sowie die Qualität der Datenanalyse. Die Gruppe der Brauchbarkeit besteht aus den folgenden Kriterien: Klarheit der Fragestellung (Definition des Untersuchungsfeldes inkl. Zieldefinition), Eindeutigkeit der Resultate, konkrete und fundierte Schlußfolgerungen, Beachtung alternativer Interpretationsmöglichkeiten. Die Ausführungen zu den Qualitätskriterien sind sehr knapp gehalten und dienen lediglich dazu, die in der Untersuchung eingesetzten Beurteilungskriterien transparent zu machen. Dies ist auch der deutlichste Unterschied zu den ‚Standards‘, die nicht – wie die Normen der Algemene Rekenkamer – dazu geschaffen wurden, eine spezifische Gruppe von Evaluationen nachträglich einer Qualitätskontrolle zu unterziehen, sondern einen deutlich weiteren Ansatz verfolgen. Auffällig ist, daß einige Kriterien der AR-Normen weitgehend mit Elementen der ‚Standards‘ übereinstimmen. Den beiden ‚Standard‘-Gruppen „Durchführbarkeit“ und „Korrektheit“ entsprechende Kriterien fehlen jedoch. Ethische Fragen werden in den AR-Normen nicht aufgegriffen.

2. Leitfaden für die Planung von Projekt- und Programmevaluationen. Bundesamt für Gesundheit, Fachbereich Evaluation, Bern 1997 (BAG-Leitfaden).

Das Bundesamt für Gesundheit (BAG) hat im Jahre 1997 den „Leitfaden für die Planung von Projekt- und Programmevaluationen“ herausgegeben. Das BAG ist auf Bundesebene für die schweizerische Gesundheitspolitik zuständig. Es läßt in diesem Zusammenhang regelmäßig Evaluationen durchführen, die sich zumeist mit Programmen oder Projekten des BAG befassen bzw. mit solchen, die durch das

BAG (mit)finanziert werden. Der Leitfaden richtet sich primär an die Mitarbeiter des BAG, die Evaluationsprojekte zu begleiten haben, und an die extern mit Evaluationen des BAG betrauten Stellen. Der BAG-Leitfaden ist mit achtzig Seiten recht umfangreich. Abgesehen von einer in die Evaluation einführenden Einleitung und verschiedenen hilfreichen Anlagen (Literatur, Terminologie, Übersicht über Evaluationsparadigmen, Typologie von Evaluationsfragen usw.) ist der BAG-Leitfaden entlang des Evaluationsprozesses strukturiert: ‚Planen einer Evaluation‘, ‚Erarbeiten, Mandatieren und Begleiten einer Evaluation‘, ‚Der Evaluationsbericht‘, ‚Verbreitung und Verwendung von Evaluationsergebnissen‘. Diese vier Kapitel enthalten neben erläuternden und erklärenden Ausführungen eine oder mehrere Checklisten, die durch das BAG aber auch durch die externen Stellen eingesetzt werden können. Mit diesen Checklisten soll sichergestellt werden, daß Evaluationen der Evaluationspolitik des BAG entsprechen. Es sind nur in wenigen Aspekten BAG-spezifische Besonderheiten vorhanden, die eine Übertragung auf andere Kontexte erschweren würden. Zumeist beziehen sich die im BAG-Leitfaden enthaltenen Normen auf allgemein in der Evaluation verbreitete Prinzipien.

Im Vergleich zu den ‚Standards‘ fällt an erster Stelle die Orientierung an den Bedürfnissen des BAG auf. Zweitens ist der BAG-Leitfaden in seinem Aufbau eng an einem idealtypischen Ablauf eines Evaluationsprozesses orientiert. Daher erhalten Aspekte des Projektmanagements im BAG-Leitfaden deutlich größeres Gewicht als bei den ‚Standards‘. Eine Übereinstimmung mit den ‚Standards‘ besteht darin, daß nicht nur Evaluatorinnen, sondern auch die Seite des Auftraggebers und weitere Beteiligte mit dem BAG-Leitfaden angesprochen werden. Der BAG-Leitfaden wählt aber dafür ein etwas anderes Vorgehen. So spricht der BAG-Leitfaden die genannten Adressatenkreise jeweils mit spezifischen Checklisten an. Die ‚Standards‘ hingegen richten sich im allgemeinen an den ganzen Adressatenkreis, ohne Zuständigkeiten oder Verantwortlichkeiten festzulegen.

3. Standards for Program Evaluation. Evaluation Research Society Standards Committee. In: New Directions for Program Evaluation, No. 15 (1982). Deutsche Übersetzung: Standards der Evaluation der Evaluation Research Society. In: Koch, Uwe/Wittmann, Werner W. (Hrsg.): Evaluationsforschung. Berlin: Springer 1990: 311-5. (ERS-Standards)

Die ERS-Standards (siehe dazu auch Widmer 1996a:11) sind durch ein Komitee der ‚Evaluation Research Society‘ (ERS) erarbeitet worden. Diese Gesellschaft stellte eine berufsständische Organisation vor allem US-amerikanischer Evaluatoreninnen dar. Im Jahre 1985 hat sie sich mit dem mitgliedermäßig kleineren ‚Evaluation Network‘ (ENet) zur ‚American Evaluation Association‘ (AEA, vgl. unten) zusammengeschlossen (vgl. Berk and Rossi 1990:110-1). Die ERS-Standards sind zwar in der Literatur der frühen achtziger Jahre relativ breit diskutiert worden, das Interesse hat aber seither deutlich abgenommen.

Die ERS-Standards setzen sich aus sechs Gruppen mit insgesamt 55 Einzelstandards zusammen. Dabei werden folgende Themen behandelt: Formulierung und Aushandlung der Evaluation, Struktur und Planung der Evaluation, Datenerhebung und -aufbereitung, Datenauswertung und -interpretation, Kommunikation und Bekanntgabe der Ergebnisse sowie deren Verwendung. Die ERS-Standards sind somit sequentiell, entlang der Phasen eines traditionellen Evaluationsprozesses gruppiert. Die in den ERS-Standards vertretene Evaluationsauffassung orientiert sich an einem objektivistischen Wissenschaftsverständnis. Naturalistische (Guba and Lincoln 1989), nützlichkeitsfokussierte (Patton 1997) oder responsive (Beywl 1988) Ansätze ebenso wie die neuere ‚empowerment evaluation‘ (Fetterman 1996) werden dadurch ausgegrenzt (vgl. Guba and Lincoln 1989: 230). Erkenntnistheoretischer Hintergrund ist ein Verständnis von wissenschaftlich absoluter Objektivität/Neutralität. Schon Lee J. Cronbach hatte dazu kritisch vermerkt, daß der Begriff ‚values‘ in den ERS-Standards kein einziges Mal verwendet wird (Cronbach 1982:53).

Diese relativ enge methodologische Ausrichtung bildet den Hauptunterschied zu den ‚Standards‘. Die ‚Standards‘ werden (im Gegensatz zu den ERS-Standards) von Vertretern alternativer Ansätze grundsätzlich akzeptiert (vgl. Guba and Lincoln 1989: 229-33, Patton 1994 und 1994a sowie 1997: 15-7, Beywl 1988: 113-23 und Fetterman 1995).

Der zweite wesentliche Unterschied, der eine direkte Folge des erstgenannten darstellt, bildet die sequentielle Struktur. Drittens sind die formellen Differenzen zu erwähnen, also einmal die doch deutlich höhere Zahl von Einzelstandards in den ERS-Standards, wie auch die bei den ERS-Standards fehlenden ergänzenden Ausführungen. Eine vierte Differenz liegt in der heutigen institutionellen Abstützung der beiden Standards. Mit den AEA-Principles aus dem Jahre 1994 (siehe unten) hat die Nachfolgeorganisation der ERS neue Qualitätsnormen erlassen. Aufgrund ihrer epistemologischen Ausrichtung wie auch wegen der fehlenden institutionellen Basis haben die ERS-Standards deutlich an Bedeutung eingebüßt. Ansonsten weisen sie aber im Vergleich zu anderen hier behandelten Qualitätsnormen verhältnismäßig viele Parallelen mit den ‚Standards‘ auf; insbesondere der Adressatenkreis und die behandelte Thematik sind sehr ähnlich gelagert.

4. Quality Assessment of Evaluation Reports: a Framework. MEANS Handbook n° 5. European Union – Regional Policy and Cohesion. Bruxelles: Commission of the European Communities, 1996.(MEANS-Handbook)

Das MEANS-Handbuch (MEANS steht für ‚Methods for Evaluating Actions of a Structural Nature‘ und bildet ein Förderungsprogramm der Europäischen Kommission) stellt eine 42-seitige Broschüre im Format A5 dar, die zur Qualität von Evaluationen beitragen soll. Das Handbuch führt insgesamt acht Beurteilungskriterien an. Diese werden im Handbuch mit Erläuterungen und explizierenden Fragen ergänzt. Die Kriterien heißen: Orientierung an den Bedürfnissen, relevante Perspektive, robuster Evaluationsplan, zuverlässige Daten, fundierte Auswertung, glaubwür-

dige Befunde, unparteiische Folgerungen und Klarheit des Berichts. Neben den Kriterien zur Qualitätsbeurteilung enthält das Handbuch auch Aussagen, wie derartige Qualitätsbeurteilungen vorgenommen werden sollen. Weiter stellt das Handbuch auch Checklisten für den Einsatz in Qualitätsbeurteilungen bereit.

Ein Vergleich des MEANS-Handbuchs mit den ‚Standards‘ ergibt: Erstens führt das MEANS-Handbuch ausnahmslos Kriterien an, die auch durch die ‚Standards‘ abgedeckt werden. In gewissen Punkten ist sogar eine verblüffende Übereinstimmung festzustellen. Die ‚Standards‘ werden denn auch explizit im Handbuch aufgeführt. Zweitens ist der Fokus des Handbuchs deutlich schmaler als jener der ‚Standards‘. Beispielsweise werden ethische Fragen kaum angesprochen. Ähnliches gilt auch für vertragliche Vereinbarungen oder das Thema Glaubwürdigkeit. Hingegen wird die Meta-Evaluation, wie in den ‚Standards‘, sehr deutlich angesprochen, da ja das Handbuch explizit zur Durchführung von Qualitätsbewertung dienen soll.

5. Guiding Principles for Program Evaluation in Ontario Health Units, hrsg. vom Ontario Ministry of Health/Ottawa-Carleton Health Department, 1997 (Ontario-Principles)

Dieses Papier ist gemeinsam vom Gesundheitsministerium der kanadischen Provinz Ontario und der Gesundheitsabteilung der Stadt Ottawa-Carleton herausgegeben. Es besticht durch seine Kürze (knapp 4000 Zeichen) und seine sehr klare, praxisnahe Sprache. Ziel ist es, die Evaluation von öffentlichen Gesundheitsprogrammen zu unterstützen.

Anschließend an kurze Definitionen von „Prinzipien“, „Programm“, „Programmevaluation“ sowie „Beteiligte & Betroffene“ folgen insgesamt elf Leitprinzipien, welche aus zwei bis max. fünf jeweils kurzen Soll-Aussagen bestehen. (Diese sind in vier Hauptabschnitte: Wann, wie, wer und warum gegliedert, die jedoch wenig Orientierung geben).

Die Ontario-Principles legen großen Wert darauf, daß die Evaluation in die Programmplanung integriert ist, was eine klare Beschreibung des Programms mit seinen Zielen sowie einen explizit identifizierten Entscheidungsbedarf erfordert. Evaluationsfragestellungen sollten zentrale Steuerungsaufgaben der Programme betreffen und unmittelbar mit den Prozeß- oder Outcome-Zielen des Programms in Zusammenhang stehen. Weitere Gesichtspunkte sind der Einsatz systematischer Methoden, die Klarheit und Präzision der Berichte wie deren rechtzeitige Verbreitung. Auch bei den beiden Prinzipien „Einbezug der Beteiligten & Betroffenen“ und „Nutzung der Evaluationsbefunde“ zeigen sich große Ähnlichkeiten zur Stoßrichtung der ‚Standards‘. Ganz ausdrücklich wird für multidisziplinäre Teams von Evaluatorinnen plädiert; ethische Gesichtspunkte werden mit einem Prinzip, das aus einem einzigen relativ undifferenzierten Satz besteht, sehr knapp angesprochen.

Dieses Papier ist ein Beispiel dafür, wie für ein bestimmtes Politikfeld auf Länderebene orientierende und handhabbare Evaluationsnormen formuliert werden können.

6. Best Practice Guidelines for Evaluation. Public Management Service bei der OECD, 1998 (PUMA-Guidelines)

Die PUMA-Guidelines sollen Mitgliedsstaaten der OECD dabei unterstützen, die Nutzung von Evaluation im Rahmen von Performance-Management-Systemen zu verbessern. Sie wenden sich vorrangig an diejenigen, die für die politische Steuerung von Evaluation verantwortlich sind (Regierungsorganisationen, Politiker und leitende Beamte). Insgesamt werden neun Richtlinien mit jeweils zwei bis fünf thesenartig gefaßten Absätzen formuliert. Sie sind in drei Hauptabschnitte gegliedert: Wie man das meiste von Evaluationen bekommt – Wie man den Rahmen für Evaluation gestaltet – Aufbau effektiver Evaluationen.

Evaluation wird nicht in erster Linie als wissenschaftlich gestützte Datenerhebungsmethode verstanden, sondern darüber hinaus als originärer methodischer Ansatz zur Bewertung von fachpolitischen Programmen und Maßnahmen.

Die summative Funktion von Evaluation steht in diesem Dokument deutlich im Vordergrund (Verbesserung von Entscheidungsprozessen, Ressourcenallokation und Verantwortlichkeit). Die Orientierung auf Entscheidungen und Entscheidende (Auftraggeber, Commissioner) ist stark ausgeprägt: „Erfolgreiche Evaluationen basieren auf der Zusammenarbeit zwischen den zentralen Beteiligten (Evaluatoren, Nutzer und Beteiligte & Betroffene) unter der Leitung eines Auftraggebers. Er plant die Evaluation, ist verantwortlich für die Definition der Zielsetzungen, der Kriterien, der Datengewinnung und der Methoden“.

Dem gegenüber werden hohe Anforderungen an den Einbezug von Beteiligten & Betroffenen formuliert: „Evaluationen ohne ‚Eignerschaft‘ [„ownership“] durch Beteiligte & Betroffene werden wahrscheinlich unwirksam sein“. Partizipative Evaluationsmethoden seien daher erforderlich, um Konsens zu schaffen und eine ‚Eignerschaft‘ für einen Prozeß des Wandels zu unterstützen. „Der Dialog mit Nutzern und Mitarbeitern verbessert das Verständnis und die Responsivität gegenüber deren Anforderungen und Prioritäten“.

Die Entwicklung einer Evaluationskultur wird als wichtige Aufgabe auch auf der Ebene supranationaler Organisationen gesehen. Dazu zählt neben dem Bekenntnis zur Stakeholder-Orientierung auch die Befürwortung der Veröffentlichung von Evaluationsergebnissen und die Betonung des Faktors Glaubwürdigkeit von Evaluationen und Evaluatorinnen. Schließlich wird auch der Stellenwert professioneller und ethischer Standards für Qualität von Evaluationen herausgestellt.

Bei zahlreichen inhaltlichen Übereinstimmungen mit den ‚Standards‘ privilegieren die PUMA-Guidelines die Nutzungsart ‚Entscheidungsvorbereitung‘ deutlich gegenüber anderen und schaffen eine erhebliche Spannung zwischen Entscheidungsorientierung und partizipativem Anspruch.

7. Prinzipien zur Durchführung von Evaluationen im Bereich der Wissenschaft. Schweizerischer Wissenschaftsrat (SWR), SWR-DOK.96-2. Bern: SWR, 14. Dezember 1995.

und

Codex Ethik Wissenschaftsevaluation, Schweizerischer Wissenschaftsrat (SWR), SWR-DOK.96-3. Bern: SWR, 14. Dezember 1995.

Der Schweizerische Wissenschaftsrat (SWR) ist das beratende Organ des Bundesrates (Regierung) für alle wissenschaftspolitischen Fragen und setzt sich hauptsächlich aus Wissenschaftlern sowie aus Verantwortlichen wissenschaftlicher, wirtschaftlicher und politischer Institutionen zusammen. In seinem Pflichtenheft erscheint als einer von vier Punkten die Evaluation der Wissenschafts- und Technologiepolitik. Der SWR gibt regelmäßig Evaluationen an externe Experten in Auftrag und führt selber Evaluationen durch. In diesem Zusammenhang verabschiedete der SWR am 14. Dezember 1995 eine Grundsutzerklärung mit dem Titel „Prinzipien zur Durchführung von Evaluationen im Bereich der Wissenschaft“ (SWR-Prinzipien). Dieses Papier im Umfang von zwei Seiten wird ergänzt durch eine Beilage „Codex Ethik Wissenschaftsevaluation“ (SWR-Codex) mit gleichem Umfang. Der Geltungsbereich der in den beiden Papieren enthaltenen Regeln soll vorerst – momentan befinden sich die Unterlagen in Überarbeitung – auf Evaluationen beschränkt bleiben, die vom SWR durchgeführt oder an denen er beteiligt ist. Die SWR-Prinzipien enthalten – verkürzt dargestellt – folgende Punkte: Evaluationen sollen primär im Dienste der Wissenschaft stehen. Die Prinzipien führen, nach Evaluationsgegenstand differenziert, spezifische Evaluationskriterien an. Weiter legen die Prinzipien fest, daß Evaluationen auf wissenschaftlichen Grundlagen zu beruhen haben. Sie regeln zudem, daß Evaluationsbestrebungen koordiniert erfolgen, daß die Evaluationsergebnisse den Betroffenen zugänglich gemacht werden und daß SWR-Evaluationen einen direkten Bezug zu Prozessen der Wissenschaftspolitik aufweisen sollen. Für ethische Fragen wird in den Prinzipien auf die Beilage verwiesen, den SWR-Codex. Dieser umfaßt aber nicht nur Festlegungen ethischer Art im engeren Sinne (wie Würde von Personen und Institutionen, Regeln der Kommunikationskultur usw.), sondern trifft über ethische Fragen hinausgehende Festlegungen. So wird fixiert, daß Evaluationspläne erstellt werden, daß Evaluationen Entscheidungsgrundlagen bereitstellen oder daß Wissenschaftsevaluationen nicht unnötigerweise Ressourcen und Kapazitäten der Forschung blockieren sollen. Zum Teil werden auch Normen, die bereits in den SWR-Prinzipien angeführt sind, im SWR-Codex wiederholt (etwa Evaluation auf wissenschaftlicher Grundlage).

Im Vergleich zu anderen hier vorgestellten Normen (inkl. den ‚Standards‘) läßt sich zum einen feststellen, daß SWR-Prinzipien und -Codex relativ wenig systematisiert erscheinen. Eine zweite Besonderheit im Vergleich mit den ‚Standards‘ bilden die relativ engen institutionellen Rahmensetzungen, in denen sich die SWR-Regelwerke bewegen. Es handelt sich explizit um Normen, die für selbstdurchgeführte Evaluationen einer Wissenschaftsinstitution Regeln aufstellen. Deshalb eig-

nen sie sich auch weniger für eine Anwendung über den durch die Institution gesetzten Rahmen hinaus, wie dies für andere, hier diskutierte Qualitätsnormen gilt.

V. Evaluationsspezifische Regelwerke, die auf Fachpersonen gerichtet sind

1. Guiding Principles for Evaluators. American Evaluation Association (AEA), Task Force on Guiding Principles for Evaluators. In: Evaluation Practice News, June 1994. Reprint in: New Directions for Program Evaluation, No. 66 (Summer 1995): 19-26. (Deutsche Übersetzung in Beywl 1996) (AEA-Principles)

Im Juni 1994 hat die ‚American Evaluation Association‘ (AEA) die ‚Guiding Principles for Evaluators‘ herausgegeben. Im folgenden Jahr erschien eine Sondernummer der Zeitschrift ‚New Directions for Program Evaluation‘, in der neben den AEA-Principles verschiedene begleitende Texte enthalten sind (Shadish et al. 1995; vgl. auch Widmer 1996a: 50-1). Die AEA-Principles umfassen folgende fünf Leitprinzipien:

1. *Systematische Untersuchung*: Evaluаторinnen führen systematische, auf Daten gestützte Untersuchungen über das jeweilige Evaluationsobjekt durch.
2. *Kompetenz*: Evaluаторinnen stellen den Beteiligten & Betroffenen professionelle Leistungen zur Verfügung.
3. *Integrität/Aufrichtigkeit*: Evaluаторinnen garantieren Aufrichtigkeit und Integrität während des gesamten Evaluationsprozesses.
4. *Achtung gegenüber den Menschen*: Evaluаторinnen respektieren die Sicherheit, die Würde und das Selbstwertgefühl der Antwortenden, Programmteilnehmenden, Auftraggeber und anderer Beteiligter & Betroffener, mit denen sie in Interaktion stehen.
5. *Verantwortung für das allgemeine und öffentliche Wohl*: Evaluаторinnen artikulieren die vielfältigen Interessen und Werte, die möglicherweise in Beziehung zum allgemeinen und öffentlichen Wohl stehen, und berücksichtigen diese in ihren Überlegungen.

Diese fünf Prinzipien werden von einer Reihe weitergehender Überlegungen begleitet, wobei diese Erläuterungen weder alle Anwendungen der AEA-Principles darstellen sollen, noch als verbindliche Regeln zu verstehen sind, die etwa als Grundlage zur Sanktionierung von Verstößen dienen könnten.

Im Vergleich der AEA-Principles mit den ‚Standards‘ fallen drei Aspekte auf (siehe dazu auch Sanders 1995 und Widmer 1996a: 50-1). Erstens ergeben sich zwischen den beiden Texten deutliche inhaltliche Übereinstimmungen. Alle in den AEA-Principles enthaltenen Ansprüche finden sich in der einen oder anderen Form

in den ‚Standards‘ wieder. Zweitens besteht ein deutlicher Unterschied in der Ausführlichkeit: Die AEA-Principles sind viel genereller angelegt, weniger spezifiziert und kürzer gefaßt als die ‚Standards‘ – dies ist beabsichtigt, sollten doch die AEA-Principles die ‚Standards‘ nicht ersetzen sondern ergänzen. Durch ihre Knappheit und durch ihren übergreifenden Charakter stellen die AEA-Principles eine willkommene Orientierungshilfe für die Praxis dar. Bei den ‚Standards‘ besteht allein wegen ihres Umfangs die Gefahr, daß der pragmatisch orientierte Praktiker sie nicht heranzieht (was allerdings durch die Kurzfassung der Standards, die vielfach in Lehrbüchern und Sammelwerken zur Evaluation reproduziert ist, kompensiert wird). Drittens richten sich die AEA-Principles direkt an die Evaluatorin. Sie stellen also Verhaltensrichtlinien für das fachliche Handeln der Evaluationsfachleute dar, im Gegensatz zu den ‚Standards‘, die sich auf Evaluationen als Prozesse der Leistungserstellung beziehen (zum Vergleich siehe etwa Covert 1995, Newman 1995: 108, Sanders 1995, Widmer 1996a: 50-1 sowie Beywl 1996).

2. Guidelines for the Ethical Conduct of Evaluations. Australasian Evaluation Society. Curtin ACT: AES, January 1998. (AES-Guidelines).

Die ‚Australasian Evaluation Society‘ (AES) bildet eine professionelle Evaluationsorganisation, deren Mitglieder vor allem aus Australien und Neuseeland stammen. Im Jahre 1996 wurde ein ‚Committee on Ethics and Standards in Evaluation‘ eingesetzt, mit der Aufgabe, Leitlinien für ein ethisch korrektes Verhalten bei Evaluationen zu erarbeiten. Die so entstandenen ‚Guidelines for the Ethical Conduct of Evaluations‘ (AES-Guidelines) wurden im Dezember 1997 vom Vorstand der AES genehmigt und traten Anfang 1998 in Kraft. Die Leitlinien sollen ständig weiterentwickelt werden.

Die AES-Guidelines sollen die Diskussion über ethische Fragen stimulieren und den Mitgliedern dabei helfen, ethische Probleme zu erkennen und sich in konkreten Situationen richtig zu verhalten. Weiter sollen die AES-Guidelines Eingang in Handbücher und Leitfäden zu Evaluationsfragen finden. Hingegen verzichtet die AES darauf, die Einhaltung der Leitlinien zu überwachen oder gar zu sanktionieren. Vielmehr sollen die Mitglieder der AES ihre ethische Verantwortung selbstverantwortlich wahrnehmen.

Die AES-Guidelines richten sich nicht nur an Evaluatorinnen, sondern auch an Auftraggeber von Evaluationen und weitere Kreise, die Evaluationen nutzen. Angesprochen sind auch Personen, die in Evaluationsfragen forschen, unterrichten oder publizieren.

Die AES-Guidelines sind unterteilt in drei Kapitel, nämlich Auftragsvergabe und Vorbereitung von Evaluationen, Durchführung von Evaluationen sowie Berichterstattung über Evaluationsergebnisse. Insgesamt umfassen die AES-Guidelines 22 Leitsätze. Inhaltlich handelt es sich dabei schwergewichtig um Aussagen, die das faire, ethisch verantwortliche, offene und aufrichtige Verhalten der an Evaluationen Beteiligten festlegen.

Im Gegensatz zu den ‚Standards‘, die in den AES-Guidelines explizit als komplementäre Normierung Erwähnung finden und vom Vorstand der AES im August 1996 gutgeheißen wurden, enthalten die AES-Guidelines keine Hinweise methodischer Art. Sie konzentrieren sich auf ethische Fragen, wobei ein relativ weites Verständnis von Ethik zugrundeliegt⁵².

Im Gegensatz zu anderen hier präsentierten Qualitätsnormen zeichnen sich die AES-Guidelines dadurch aus, daß sie mehrere an Evaluationen beteiligte Personengruppen ansprechen. Hier ist auch eine wesentliche Übereinstimmung mit den ‚Standards‘ festzuhalten, während Unterschiede besonders für die abgedeckte Thematik zu erkennen sind – die ‚Standards‘ beschränken sich nicht auf ethische Fragen. Auch der Umfang der AES-Guidelines ist deutlich geringer. Begleitende Texte, Erläuterungen, Fallbeispiele oder Hinweise auf häufige Fehler sind in den AES-Guidelines nicht enthalten.

3. *CES Guidelines for Ethical Conduct. Canadian Evaluation Society. o.O., o.J. (CES-Leitlinien)*

Die CES-Leitlinien gehören zu den kürzesten hier vorgestellten Regelwerken. Sie umfassen nur drei Gruppen, nämlich: Kompetenz, Integrität und Verantwortlichkeit, die jeweils drei bis vier Einzelpunkte enthalten. Unter den Begriff der *Kompetenz* fällt die Verwendung systematischer und angemessener Untersuchungsmethoden, die Verfügbarkeit ausreichenden inhaltlichen Wissens sowie das Bemühen der Evaluatorin, ihr methodisches und praktisches Können kontinuierlich zu verbessern. Zur *Integrität* einer Evaluatorin zählen: Transparenz hinsichtlich des eigenen Wissens und der eigenen Fähigkeiten, die Verpflichtung zur Deklaration auftretender Interessenkonflikte, kulturelle und soziale Sensitivität und die Einhaltung der eingegangenen vertraglichen Verpflichtungen. Zur *Verantwortlichkeit* werden aufgeführt: dem Klienten durch entsprechende Informationen die Entscheidungsfindung zur angemessenen Evaluationsstrategie und -methodik (inkl. deren Grenzen) zu erleichtern; für eine klare, angemessene und faire mündliche und schriftliche Präsentation der Evaluationsergebnisse zu sorgen; Evaluationen mit einem vorteilhaften Kosten-Nutzen-Verhältnis durchzuführen sowie Einhaltung terminlicher Vorgaben, auf die man sich geeinigt hat. Der Gegenstandsbereich der CES-Guidelines betrifft damit unter anderem die Methodik, das Verhältnis zu Auftraggebern und Beteiligten & Betroffenen und die Berichterstattung. Er ist sehr weit und sehr allgemein gefaßt.

Im Vergleich zu den ‚Standards‘ fällt neben der Kürze auf, daß die CES-Guidelines Normen enthalten, die in mehr oder weniger identischer Form auch in den ‚Standards‘ enthalten sind. Die ‚Standards‘ sind aber umfangreicher, detaillierter und spezifischer. Ein weiterer wichtiger Unterschied besteht darin, daß sich die

52 So enthalten die AES-Leitlinien auch eine Richtlinie, die methodische Strenge in Design, Datenerhebung und -auswertung einfordert.

CES-Guidelines explizit und ausschließlich an die Evaluatorinnen als Handelnde wenden.

VI. Resümee

Wie die vorangegangene – bei weitem nicht vollständige – Übersicht deutlich macht, werden Regelwerke zur Gestaltung fachlicher Praxis in Evaluation und angrenzenden Feldern seit Beginn der neunziger Jahre verstärkt durch Fachvereinigungen, teils auch gemeinsame Ausschüsse verschiedener Fachvereinigungen, erarbeitet und veröffentlicht. Dies läßt auf einen hohen Bedarf nach orientierenden Grundlagen schließen, der insbesondere seitens der Fachleute besteht, aber auch – je nach Anwendungsnähe der Profession in unterschiedlichem Maße – seitens der Abnehmer solcher Leistungen und der interessierten Öffentlichkeit. Nicht zuletzt kommt darin auch die hohe Dynamik (und Expansion) zum Ausdruck, die in den letzten Jahren in den untersuchten Professionsfeldern vorliegt.

Der Vergleich zeigt, daß sich die Regelsetzungsprozesse und deren Produkte nach inhaltlichen und formalen Merkmalen deutlich unterscheiden, je nach beabsichtigtem Anwendungsfeld und professionellem Selbstverständnis der regelsetzenden Organisation(en). Wichtige Unterschiede und Übereinstimmungen sollen abschließend mit Bezug zu den ‚Standards‘ angesprochen werden.

1. Inhaltliche Schwerpunkte

Die verschiedenen Regelwerke setzen hinsichtlich der zu normierenden Tatbestände deutlich unterschiedliche Schwerpunkte:

1. Ethische Haltung der Fachkräfte: AEA-Principles, AES-Principles, APSA-Guide, CES-Guidelines, DGPs/BDP-Richtlinien, DGS/BDS-Kodex (inkl. der darauf basierenden Kodices der DVPW und der SVPW), DGSv-Berufsordnung;
2. Ethische Haltung der Fachkräfte mit starker Parteinahme für Benachteiligte der Gesellschaft: DBSH-Prinzipien, DGVT-Richtlinien;
3. Interaktion mit Beteiligten: AERA-Standards, CMC-Kodex, SMS-Normen/Swiss Interview-Richtlinien;
4. Ausbildung der Fachpersonen: DGSv-Standards;
5. Methodik und Vorgehensweise: AR-Normen, BAG-Leitfaden, CDEP-Standards, DFG-Kriterien, ERS-Standards, MEANS-Handbuch;
6. Nützlichkeit/Wirkung der professionellen Praxis: dv-Empfehlungen, GA-Standards, Ontario-Principles, PUMA-Guidelines, SWR-Prinzipien.

Zwischen der thematischen Schwerpunktsetzung und der Bezeichnung des jeweiligen Regelwerkes scheinen keine systematischen Zusammenhänge zu bestehen. Die Bezeichnungen Prinzipien, Leitlinien, Richtlinien, Kodex, Berufsordnung, Stan-

dards, Normen, Leitfaden, Kriterien, Handbuch, Empfehlungen gehen offensichtlich auf die Genese des Regelwerkes in seinem jeweiligen organisatorischen und/oder fachpolitischen Kontext zurück. Prinzipien oder Leitlinien werden öfter für auf Fachpersonen angewandte Regelwerke genutzt, Standards öfter für Leistungsprozesse (wie auch den Ausbildungsprozeß von Fachpersonen – DGSv-Standards) – aber es gibt auch Gegenbeispiele (PUMA-Guidelines).

Die ‚Standards‘ lassen sich keiner der aufgeführten thematischen Ausrichtungen eindeutig zuordnen. Zwar gibt es auch bei anderen Regelwerken Zuordnungsschwierigkeiten bzw. thematische Überschneidungen, die ‚Standards‘ sind jedoch offensichtlich so plazierte, daß sie gleichzeitig mehrere thematische Schwerpunkte berühren. Die ethischen Aspekte (siehe oben, Punkt 1) werden in den ‚Standards‘ schwergewichtig unter dem Begriff „Korrektheit“ angesprochen. In dieser Gruppe der ‚Standards‘ finden sich auch Ausführungen zur Gestaltung der Interaktion zwischen den Beteiligten (Punkt 3). Die Methodik und die Vorgehensweise (Punkt 5) wird unter „Genauigkeit“ subsumiert. Der Nutzen und die Wirkungen der professionellen Praxis (Punkt 6) wird hauptsächlich in der Gruppe der Nützlichkeits-‚Standards‘ thematisiert und stellt – unterstrichen durch ihre Stellung zu Beginn der insgesamt 30 Einzelstandards – einen Schwerpunkt dieses Regelwerkes dar. Ähnlich den CDEP-Standards sprechen die ‚Standards‘ die spezifischen Belange benachteiligter Gruppen lediglich kurz und beschränkt auf die Perspektive ethnischer (insbesondere nicht-muttersprachlich englischer) Minoritäten an (Punkt 2).⁵³ Die ‚Standards‘ plädieren vielmehr für einen ausgewogenen und fairen Ansatz. Fast ganz außer Acht gelassen wird in den ‚Standards‘ das Thema der Ausbildung für die professionelle Praxis (Punkt 4); allerdings empfiehlt das Joint Committee die Nutzung der ‚Standards‘ als eine Art Lehrtext für Ausbildungszwecke (vgl. das Kapitel „Anwendung der Standards, S. 31).

2. Qualitätsträger

Wer oder was trägt die durch die Regelwerke kodifizierten, als wünschenswert ausgewiesenen Qualitätsmerkmale? Sollen es die fachlichen Leistungsprozesse (u.a. die *Evaluationen*) sein oder die handelnden *Personen* (u.a. die *Evaluatorinnen*)?

Wichtigstes auf Evaluationsfachkräfte bezogenes Regelwerk sind die *Guiding Principles for Evaluators* (AEA-Principles) von 1995, außerdem die *Guidelines for the Ethical Conduct of Evaluations* der Australasian Evaluation Society von 1998 (AES-Guidelines), die sich zusätzlich insbesondere an Auftraggeber von Evaluationen richten. Als deutschsprachiges Regelset formuliert der DGS/BDK-Kodex ähnliche, auch für Evaluatorinnen sinngemäß anwendbare Anforderungen, während die DGPs/BDP-Richtlinien oder die DBSH-Prinzipien deutlich weniger unmittelbar übertragbare inhaltliche Bezüge zu einer guten Praxis von Evaluatorinnen aufweisen, bedingt durch die andere Handlungstypik der Berufsgruppen Psychologen bzw.

53 Vgl. insbesondere die Ausführungen zu den Standards N1, N3, K3 oder K6.

Sozialarbeiter (z.B. klinische Arbeit, Einzelfall- oder Gruppenansatz in der sozialen Arbeit).

Bei den auf Evaluationen als Leistungsprozesse gerichteten Normen erhalten *The Program Evaluation Standards des Joint Committee on Standards for Educational Evaluation aus dem Jahr 1994* („Standards“) die größte Beachtung. Viele andere Regelwerke dieser Kategorien nehmen explizit (z.B. Ontario-Principles) oder implizit (z.B. MEANS-Handbook) auf die „Standards“ Bezug. Im deutschsprachigen Raum findet sich kein vergleichbares oder zumindest analoges Regelset. Eine gewisse Ausnahme stellt der BAG-Leitfaden dar, der – wie der Name schon sagt – eher (rezeptartig) handlungsanleitend als (normativ) orientierend ist. Er ergänzt damit die „Standards“ um pragmatische Hinweise zum Evaluationsmanagement.

3. Regelsetzer

Welche Organisation oder welches Gremium stellt das Regelwerk auf?

Im Falle der AEA-Principles und der AES-Guidelines oder auch der *Guidelines for Ethical Conduct* der Canadian Evaluation Society sind es Berufsverbände, in denen überwiegend Evaluatoren und Evaluatorinnen Mitglieder sind. Typischerweise werden Regelwerke, die sich auf Personen als Qualitätsträger beziehen, eigenständig durch wissenschaftliche Fachorganisationen bzw. Berufsverbände erarbeitet. Im deutschsprachigen Raum tragen sie gelegentlich die Bezeichnung „Berufsordnung“, wie z.B. im Falle der *Berufsordnung der Deutschen Gesellschaft für Supervision e.V.* von 1998. Gelegentlich erstellen solche Professionsverbände auch Regelwerke, die sich auf professionelle/wissenschaftliche Leistungsprozesse beziehen, z.B. die *Standards der Schweizerischen Evaluationsgesellschaft* (SEVAL-Standards; Widmer/Landert/Bachmann 1999).

Um einen anderen Typus von regelsetzenden Einrichtungen handelt es sich bei Ministerien oder regierungsnahen Organisationen. Ein Beispiel sind die *Guiding Principles for Program Evaluation in Ontario Health Units*. Ein anderes Beispiel – nun auf internationaler Ebene – sind die *Best Practice Guidelines for Evaluation* aus dem Jahr 1998, erarbeitet vom Public Management Service bei der OECD (PUMA-Guidelines).

Um eine Mischung zwischen Berufsverbänden und Vereinigungen von Vertretern öffentlicher Einrichtungen und Behörden handelt es sich beim Joint Committee on Standards for Educational Evaluation, dem gegenwärtig 16 Organisationen, darunter überwiegend Berufsorganisationen aber auch Zusammenschlüsse von Funktionsträgern aus Behörden gehören (vgl. die Übersicht 3 im Beitrag „Die Übertragbarkeit der Evaluationsstandards ...“, S. 250). Es fällt auf, daß Anbieter von Evaluationen und Nachfrager (insbesondere aus dem Schulsektor) hier gemeinsam ein Regelwerk erarbeitet und verabschiedet haben.

Die Frage, wie *interdisziplinär* ein Regelwerk angelegt ist, hängt eng zusammen mit der regelsetzenden Instanz. Während disziplinäre, wissenschaftliche Ver-

einigungen (wie APSA, DGPs oder DGS) Regelwerke ausgesprochen fachgebunden auslegen (also z.B. politikwissenschaftlich, psychologisch oder soziologisch), trifft dies für andere Regelsetzer weniger zu. Öffentliche Stellen, welche als Auftraggeber fungieren oder die professionelle Praxis intra muros normieren (wie etwa AR, BAG, PUMA, SWR, weniger ausgeprägt das GAO), konzentrieren sich zwar auf die teils auch einzelfachlich basierte Aufgabenwahrnehmung, lassen aber innerhalb dieser Eingrenzung Multidisziplinärität zu. Eine stark multi- oder transdisziplinäre Ausrichtung wiederum weisen jene Regelwerke auf, die sich aus betont interdisziplinären Tätigkeiten heraus entwickelt haben oder die als eher neuere Professionen mit weniger tradiertem Selbstverständnis zu beschreiben sind (wie etwa AEA, AES, CES, DBSH, DGSv).

4. Sanktionsgehalt

Legen die Regelwerke Selbstverpflichtung auf freiwilliger Basis nahe oder sehen sie implizite oder explizite Sanktionsmöglichkeiten bei Verstößen oder Nicht-Beachtung vor?

Die meisten von uns ausgewerteten Regelwerke sehen keine oder zumindest keine formalen Sanktionen vor. Dies gilt z.B. für die ‚Standards‘, die als Empfehlungen und Leitlinien aufgefaßt werden sollen und in keiner Weise mit Prüfverfahren oder gar Sanktionen verknüpft sind. Es gibt allerdings einen ersten Bericht darüber, daß die ‚Standards‘ im Rahmen einer Zivilklage herangezogen wurden, in der es um die Weigerung eines Auftraggebers ging, eine aus seiner Sicht schlecht durchgeführte Evaluation vollständig zu bezahlen. Indirekte Sanktionsmechanismen enthalten auch die von öffentlichen Dienststellen herausgegebenen Regelwerke – sie werden sicher einen wichtigen Bezugspunkt bei der Beurteilung angelieferter Evaluationsleistungen darstellen, die wiederum bei der erneuten Vergabe von Evaluationsaufträgen herangezogen werden können. Direkter sanktionieren die Berufsverbände, z.B. indem Sie eine Art Ehrengericht bei Verdacht und Meldung von Verstößen vorsehen (z.B. der 1993 veröffentlichte gemeinsame *Ethik-Kodex der Deutschen Gesellschaft für Soziologie und des Berufsverbandes Deutscher Soziologen* (DGS/BDS-Kodex). Weitgehende Konsequenzen kann der Spruch des Ehrengerichtes auf Basis der *Ethischen Richtlinien der Deutschen Gesellschaft für Psychologie e.V. und des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V.* von 1998 (DGPs/BDP-Richtlinien) haben: Der Ausschluß aus dem Berufsverband kann in diesem Fall stark verbandlich regulierter Berufstätigkeit (ähnlich wie bei den Kassenärzten) zur Beeinträchtigung wenn nicht Gefährdung der wirtschaftlichen Existenz des (dann ausgeschlossenen) Mitglieds führen. Die DGSv-Standards oder die SMS-Normen/Swiss-Interview-Richtlinien regulieren eine Art verbandlichen Zusammenschluß von wirtschaftlich tätigen Ausbildungsinstituten bzw. Befragungsunternehmen und haben damit ebenfalls hohe Verbindlichkeit und recht starken Sanktionsgehalt.

5. Umfang

Wie ausführlich und umfangreich muß ein Regelwerk sein? Wie sollte es aufgebaut sein?

Der Umfang der von uns ausgewerteten Regelwerke schwankt zwischen ca. ein bis drei Druckseiten (z.B. CES-Guidelines oder Ontario-Principles) und ca. 240 Druckseiten („Standards“). Die meisten Regelwerke kommen mit ca. vier bis zwölf Druckseiten aus. Die „Standards“ verfügen über eine ca. dreiseitige Kurzfassung – dies macht sie wiederum übersichtlich und praktikabel. Bei vergleichbarem Gesamtumfang enthalten die CDEP-Standards 180 Einzelstandards, während die „Standards“ auf die Zahl von dreißig kommen. Je nachdem, ob ein Regelwerk eher Leitprinzipien für (ethisch) korrektes Handeln formuliert oder ob es differenzierte Bezugspunkte z.B. für Leistungsprozesse auf dem Stand der Zeit bieten will (z.B. CDEP-Standards und „Standards“), ist ein kürzerer oder längerer Umfang angeraten. Welcher Umfang, welcher Grad von Strukturierung oder auch Präzision in der regulierenden Aussage wünschenswert ist, um eine gute Fachpraxis vorzubereiten und fortlaufend zu orientieren, dies sind Fragen, die vielleicht in Zukunft vergleichende Evaluationen von Regelwerken zu beantworten versuchen können.

Angesichts der Vielfalt der hier präsentierten (von den nicht-vorgestellten gar nicht zu sprechen) Regelwerke könnte die Frage gestellt werden, ob nicht eine Vereinheitlichung der Regelwerke wünschenswert ist. Wir sind der Auffassung, daß eine Vereinheitlichung, etwa in der Form eines global gültigen Regelwerkes zur Qualität von Evaluatoreninnen und Evaluationen wenig erstrebenswert ist. Dies aus folgenden Gründen: Erstens spielt die Akzeptanz der Regelwerke eine große Rolle. Regelwerke die auf wenig Akzeptanz, ja wenig Beachtung stoßen, können ihre Ziele nicht erreichen. Akzeptanz kann insbesondere bei der erstmaligen Erarbeitung und dem Beschließen von Regelsets geschaffen werden. Die Ausführungen von James Sanders (vgl. oben, S. 237ff.) zur Weiterentwicklung der „Standards“ weisen darauf hin, daß Akzeptanz durch periodische Revision immer wieder hergestellt werden sollte. Zweitens dürfen die kulturellen und sprachlichen Differenzen nicht unterschätzt werden, die z.B. zwischen den USA und Westeuropa, aber auch zwischen den europäischen Staaten und Kulturräumen bestehen. Mit der vorliegenden Übersetzung der amerikanischen „Standards“ wird ein Versuch zu einer *sprachlichen* Übertragung vorgenommen (zur Übertragbarkeit der „Standards“ vergleiche den Beitrag „Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder“ in diesem Band). Welche Resonanz diese „Standards“ finden und welche Wege ihrer Nutzung, Anpassung oder Integration eingeschlagen werden, dies werden wir mit Spannung im Laufe des kommenden Jahrzehnts verfolgen. Drittens weisen die jeweiligen Anwendungsfelder für Regelwerke prägende Besonderheiten auf, die berücksichtigt werden müssen, um eine gute Praxis sicherzustellen. Die Regelwerke müssen darauf ausgerichtet werden, so spezifisch wie möglich und so allgemein wie nötig formuliert zu werden. Damit wird das bestehende Bedürfnis nach Orientierung am besten befriedigt. Ein globales Regelwerk hingegen müßte dem Anspruch einer hohen Allge-

meinheit nachkommen, was die darin formulierten Kriterien für die praktische Anwendung schlecht umsetzbar machen würde.

Unseres Erachtens sollte bei der Formulierung künftiger Regelwerke für Evaluationen und Evaluatoreninnen nicht auf den großen, in bereits bestehende Regelwerke eingeflossenen Erfahrungsschatz verzichtet werden.

Tabelle 2: Übersicht über evaluationsunspezifische Normen, die auf die Leistungsprozesse ausgerichtet sind

	CDEP-Standards	DFG-Kriterien	DGSv-Standards	dv-Empfehlungen	GA-Standards	SMS-Normen/ Swiss Interview- Richtlinien
Erstfassung	1954	1999	1990	1992	1972	k.A./1991
Neueste Fassung	1985	1999	1998	1992	1994	k.A./1993
Umfang in Zeichen	250.000	125.000-500.000	4.000	18.000	ca. 150.000	ca. 8.500/ca. 2.600
Hauptabschnitte	Standards für (I.) die Testkonstruktion und -evaluation, (II.) eine fachlich kompetente Testanwendung, (III.) für besondere Anwendungen, (IV.) für Vorgehensweisen	Methoden der Umfrageforschung (Stichproben, Fragebogen ...)/ihre Qualitätsstandards, Archivierung von Daten, Ausbildung, Rezeption der Ergebnisse in Massenmedien, Datenschutz	Eingangsvoraussetzungen/Abschlussvoraussetzungen für eine Supervisionsausbildung, Konzeptionelle Voraussetzungen	Zielsetzung der Empfehlungen, Begriffsbestimmungen und Wirkweise von Supervision, Zur Organisation von Supervision, Finanzierung	Typen von Regierungs-Audits, Generelle Standards, Standards für die Feldarbeit bei Finanz-Audits, Standards für die Berichterstattung bei Finanz-Audits, Standards für die Feldarbeit bei Leistungs-Audits, Standards für die Berichterstattung bei Leistungs-Audits	Pflichten des Forschers gegenüber Auftraggebern/Befragten, Offenlegung, Berichterstattung/Präsentation von Ergebnissen/ Stichprobengröße, Art d. Umfrage, Obligatorische Informationen, Selbstbeschränkung, Verpflichtung der Institute
Inhaltlicher Schwerpunkt	Methodische Grundlagen für Fachdiskussion zu/ Evaluation von Tests, Testanwendungen, Testeffekten	Technische Anforderungen an Methoden und Erhebungsprozesse	Minimalanforderungen an Teilnehmende, Abschlüsse und Konzepte in Supervisionausbildung	Klärung von Leistungsschwerpunkten verschiedener Supervisionsformen	Sicherstellung von gültigen Informationen zu Finanzen und Leistungen der USA (oder des am Bundesstaates)	Beziehungsgestaltung Forscher – Befragte – Auftraggeber/Berichtsanforderungen
Qualitätsträger	Tests, Testanwendungen	Umfragestudien, -forscher	Supervisions-Ausbildungen	Supervisionen	Audits/Überprüfungen	Umfragen
Zuständiges Gremium	Committee to Develop Standards for Educational and Psychological Testing	Deutsche Forschungsgemeinschaft	Standardüberprüfungsausschuss	Vorstand des Deutschen Vereins (und Ad-Hoc-Gruppe)	United States General Accounting Office	Vereinigung Schweiz. Markt- und Meinungsforschungsinstitute
Sanktionsgehalt	gering	gering	hoch	gering	hoch	Mittel
Interdisziplinarität	mittel	mittel	hoch	mittel	gering	Mittel
Internet-Adresse http://www.	/.	/.	dgsv.de/	/.	gao.gov/	/.

Tabelle 3: Übersicht über evaluationsunspezifische Normen, die auf Fachpersonen gerichtet sind

Kurzname	AERA-Standards	AFSA-Guide	CMC-Kodex	DBSH-Prinzipien	DGP/BDP-Richtlinien	DGS/BDK-Kodex	DGSv-Berufsordnung	DGVt-Richtlinien
Erstfassung	1992	1968	1993	1997	1998	1993	1997	1997
Neueste Fassung	1992	1998	1993	1997	1998	1993	1997	1997
Umfang in Zeichen	19.000	60.000	9.000	13.000	40.000	12.000	2.500	14.000
Hauptabschnitte	Verantwortlichkeiten gegenüber dem Feld, Untersuchungspopulation, Institutionen in Erziehung/Bildung, Urheberrecht, Herausgabe, Kritik und Einschätzung von Forschung, Geldgeber, Politik und andere Nutzer, Studierende und studentische Hilfskräfte	Ethik, Verhalten (Akad. Freiheit, Verantwortung gegenüber Studierenden, Politische Aktivitäten, bzgl. Publikationen/Anstellungen/Beförderungen, Sexuelle Belästigung), Zuwendungsabhandlungen	Vertraulichkeit, Unrealistische Erwartungen, Finanzielle Interessen, (unverträgliche) Aufträge, Verhandlungen mit Klienten, Rekrutierung, Vorgehen, Andere Berater	Allgemeine Grundsätze, Verhalten gegenüber Klienten/Berufskollegen/anderen Berufen/anderen Berufsgruppen/Arbeitgebern/Öffentlichkeit	Präambel (Leitbild); tätigkeitsfeldunabhängige Bestimmungen zu Auszubildenden, Psychologinnen, Psychologen in Wissenschaft und Lehre, (Verstöße)	Präambel, Forderung (Integrität, Objektivität, Rechte der Untersuchten), Publikationen, Begutachtung, Umgang mit Studierenden, Mitarbeitern, Kollegen, Ethik-Kommission (Zusammenfassung, Aufgabebefugnisse, Sanktionen)	Allgemeine Berufspflichten, Berufsbezeichnung, Kontrakte, Vertrauensschutz, Kollektives Verhalten, Fortbildung und Kontrolle, Werbung, Konfliktmoderation	Haltung, Beziehung, Klarheit/Transparenz, Qualitätsicherung, Dialog der Lebenswelten und Kulturen, Gesellschaftspolitische Verantwortung
Inhaltlicher Schwerpunkt	Beziehungsgestaltung zu Personen/Organisationen	Disziplinäre Ethik (Lehr- und Publikationsfähigkeit)	Redliches und professionelles Geschäftsgebahren (Berufsbild Unternehmensberater)	Forderung nach Positionsnahme für die (sozial oft benachteiligten) Klientel	Ethische Verantwortung des berufstätigen Psychologen („in allen Lebenssituationen“)	Disziplinäre sowie Berufsethik	Rahmenregeln für wertebasiertes und professionelles berufliches Handeln	Gegenüber Klienten/Gesellschaft anzureichende ethische Haltungen/Handlungen
Qualitätsträger	Bildungsforscher	Politikwissenschaftler	Berater	Sozialarbeiter	Psychologen	Soziologen	Supervisoren	Fachkräfte im psychosozialen Bereich

Kurzname	AERA-Standards	AFSA-Guide	CMC-Kodex	DBSH-Prinzipien	DGFs/BDP-Richtlinien	DGS/BDS-Kodex	DGSv-Beurufung	DGYT-Richtlinien
Zuständiges Gremium	American Educational Research Association (keine Überprüfung)	Ethikkommission der APSA	Zertifizierungskommission d. Schweiz. Vereinigung der Unternehmer (ASCO)	Kommission	Mitgliederver-sammlung DGFs/Delegiertenver-sammlung BDP	Ethik-Kommission aus DGS und BDS	(nicht genannt)	(nicht genannt)
Sanktionsgehalt	Gering	mittel	hoch	Gering	Hoch	Mittel	mittel	gering
Interdisziplinarität	Mittel	gering	gering	Hoch	Gering	gering	hoch	hoch
Internet-Adresse http://www.aera.net/resource/ethics.html	aera.net/resource/ethics.html	apsanet.org/PS/ethics.html	asco.ch/	dbsh.de	dgfs.de/ gesellschaft/mittelungen/ethikrichtlinien.html	userpage.fu-berlin.de/~ifs/bds/ethikod.html	dgsv.de	dgyt.de/dgyt/ethik.htm

Tabelle 4: Übersicht über evaluationsspezifische Normen, die auf Leistungsprozesse ausgerichtet sind

Kurzname	AR-Normen	BAG-Leitfaden	ERS-Standards	MEANS-Handbuch	Ontario-Principles	PUMA-Guidelines	SWR-Prinzipien
Erstfassung	1996	1997	1982	1996	1997	1998	1995
Neueste Fassung	1996	1997	1982	1996	1997	1998	1995
Umfang in Zeichen	2.700	115.000	25.000	ca. 35.000	4.000	9.000	10.000
Hauptabschnitte	Technische Qualität, Brauchbarkeit	Planen, Erarbeiten, Mandatieren und Begleiten einer Evaluation, Evaluationsbericht, Verbreitung und Verwendung von Untersuchungsergebnissen	Formulieren und Aushandeln, Struktur und Design, Datenerhebung und -aufbereitung, Datenauswertung und Interpretation, Kommunikation und Verbreitung, Nutzung der Resultate	Bedürfnisorientierung, Relevanz, Untersuchungsanfrage, Zuverlässige Daten, Fundierte Auswertung, Glaubwürdige Ergebnisse, Ausgewogene Folgerungen, Klarheit des Berichts	Definitionen, Integration von Programmplanung und Evaluation, Programme-schreibung, Ethik, Methoden, Berichte, Multidisziplinarität, Beteiligung, Nutzung	(s. Text); Definition, Zielsetzung von Evaluation, Beteiligten-Identifizierung und -Einbezug, Evaluationskultur, Kosteneffektivität, strategische Evaluationsplanung, Glaubwürdigkeit, Verknüpfung mit Entscheidungen	Geltungsbereich, Prinzipien (Dienstleistung zuhauenden Wissenschaft, Bewertungskriterien, Wissenschaftlichkeit, Koordinations, Zugang der Betroffenen, Ethische Grundlagen)
Inhaltlicher Schwerpunkt	methodische Strenge/ Nützlichkeit	Evaluationsmanagement	Wissenschaftlichkeit des Evaluationsprozesses	Methodik und – beschränkt – Nützlichkeit	Integration von Programmplanung und Evaluation	Entscheidungsvorbereitung durch Evaluation	Nützlichkeit für die Wissenschaft, Wissenschaftlichkeit und Ethik
Qualitätsträger	Evaluationen	Evaluationen	Evaluationen	Evaluationen	Evaluationen	Evaluationen	Evaluationen
Zuständiges Gremium	Algemene Rekenkamer, NL	Fachbereich Evaluation, Bundesamt für Gesundheit, CH	ERS Standards Committee (besteht heute nicht mehr)	Centre for European Evaluation Expertise, Lyon F (im Auftrag DG XVI/B2)	Ad-Hoc-Beratungsgruppe	Public Management Committee der OECD	Schweizerischer Wissenschaftsrat
Sanktionsgehalt	gering	gering	gering	gering	gering	gering	gering
Interdisziplinarität	mittel	mittel	mittel	hoch	hoch	mittel	mittel
Internet-Adresse http://www.	–	Admin.ch/bag/ (nur Organisation)	–	–	uoftawa.ca/academic/med/epid/principale.htm	oecd.org/puma/mgmtres/pac/pubs/evaluation.n.pdf	admin.ch/swf/ (nur Organisation)

Tabelle 5: Übersicht über evaluationsspezifische Normen, die auf Fachpersonen gerichtet sind

Kurzname	AEA-Principles	AES-Guidelines	CES-Guidelines
Erstfassung	1994	1998	k.A.
Neueste Fassung	1994	1998	k.A.
Umfang in Zeichen	16.000	13.000	1.700
Hauptabschnitte	Systematische Untersuchung, Kompetenz, Integrität und Aufrichtigkeit, Achtung gegenüber den Menschen, Verantwortung für das allgemeine und öffentliche Wohl	Auftragsvergabe und Vorbereitung von Evaluationen, Durchführung von Evaluationen, Berichterstattung über Evaluationsergebnisse	Kompetenz, Integrität, Verantwortlichkeit
Inhaltlicher Schwerpunkt	Ethisch wünschbare Evaluationspraxis	Ethisches Verhalten	Ethisches Verhalten
Qualitätsträger	Evaluatorinnen	Evaluatorinnen und Auftraggeber	Evaluatorinnen
Zuständiges Gremium	Task Force on Guiding Principles for Evaluators der AEA	Ethics Task Force der AES	CES
Sanktionsgehalt	gering	gering	gering
Interdisziplinarität	hoch	hoch	hoch
Internet-Adresse http://www.eval.org/AEADocuments/documents.htm	eval.org/AEADocuments/documents.htm	203.32.109.1/aes/guidelines.htm	unites.uqam.ca/ces/ethics.html

* Die Hyperlinks werden ständig auf unserer Internet-Seite <http://www.univation.org> aktualisiert.